Linear Normal Mode Analysis of Baroclinic Instability in a Meridional Channel

by

Alison Walker
B.Sc., University of Sydney, 1995

Submitted to the Joint Program in Physical Oceanography
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

March 2001

Linear Normal Mode Analysis of Baroclinic Instability in a Meridional Channel

by

Alison Walker

B.Sc., University of Sydney, 1995

Submitted to the Joint Program in Physical Oceanography
on 13th March, 2001, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

Numerical solutions of the unstable, growing modes are found for the two-layer inviscid quasigeostrophic equations in a meridional channel. A steady mean flow in the N-S direction is imposed in the upper layer, and it is assumed that changes in planetary vorticity following this mean flow are balanced by the input of vorticity from an imposed wind stress curl. Thus in the two-layer system, the vertical shear, in thermal wind balance, is associated with an interface slope which provides a gradient of potential vorticity (PV) in the x-direction, of equal magnitude and opposite sign in the two layers. In the y-direction the PV gradient has the value of planetary beta, β, in both layers.

The unstable modes of this system exhibit a boundary-layer structure across the channel. They are intensified in the west. The growth rates of the unstable modes are of the same order as the zonal case, however the range of wavenumber and shear for which instability is possible is larger. Established cutoff criteria for the equal-layer zonal case are not applicable, and no analogous criteria has yet been found. Growing modes are found even for very weakly sheared flows, and this suggests that baroclinic instability may represent a more significant source of mid-ocean eddy energy than previously believed.

Thesis Supervisor: Joseph Pedlosky
Title: Senior Scientist, Henry, L., and Grace Doherty Oceanographer
Woods Hole Oceanographic Institute

Acknowledgments

I am very grateful of the patience and support of Joseph Pedlosky. The topic of this thesis came from his suggestion box and I have enjoyed it immensely.

This research was supported by the National Science Foundation under grant 9901654.

Contents

List of Figures

Chapter 1

Introduction

This thesis explains the initial linear evolution of perturbations on a bounded meridional flow which is sheared in the vertical.

Chapter one describes the relevant historical discoveries in baroclinic instability and the motivation for this thesis.

Chapter two introduces the governing equations, the physical setup and the method of solving the equations that leads to the calculation of the normal growing modes of the system. This chapter also includes an analysis of some of the apriori properties of the equations.

Chapter three comprises all the numerical results.

Chapter four begins by considering the role of ∇PV vectors, followed by a summary and discussion of the results. The chapter ends with a concise one-page summary of the essential contribution of the thesis.

Appendix A is a brief summary of the results of the two-layer zonal case for ease of comparison.

Appendix B documents numerical details and code.

1.1 What is baroclinic instability and why is it important ?

Instability of any system is established if a random perturbation made to that system takes the system away from its original state, ie. the perturbation grows with respect to that state.

Baroclinic instability is the instability of a mean state that is in thermal-wind balance, ie. vertical variations in the mean horizontal velocity in a rotating coordinate system are in balance with horizontal gradient of density. The horizontal variations in density set up horizontal variations in potential vorticity (PV). It is the requirement that PV be conserved on a fluid parcel can leads to the growth of perturbations.

Here I offer only a brief illustration of the baroclinic instability mechanism. Cushman-Roisin[1994] gives an excellent physically intuitive introduction with the help of schematics such as the one pictured below. Figure 1-1 is a 3D schematic of the

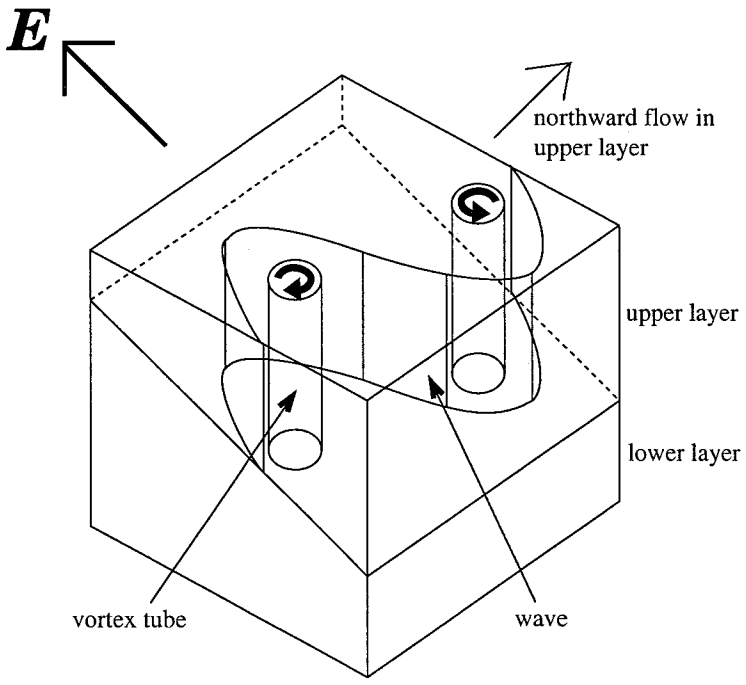

Figure 1-1: A 3D view of baroclinic instability. Only the upper layer perturbation wave is shown for clarity. The flexibility of the interface leads to the upper layer wave inducing relative vorticity, depicted as vortex tubes, of the same sign in both layers. The torque on the fluid surrounding the point vortices is such to cause growth of the perturbation waveform.

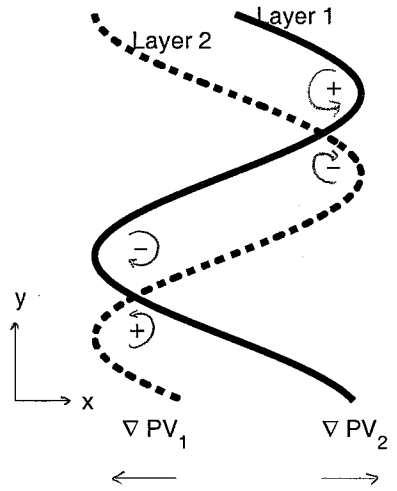

Figure 1-2: Simple schematic of how the phase shift in baroclinic instability leads to growth of the waves. The plus and minus signs indicate the sense of relative vorticity acquired by the fluid as it moves to this location. The vortices induced in one layer apply a torque on the fluid elements of the other layer, increasing the amplitude in both layers.

baroclinic instability mechanism on an f-plane. The perturbation wave is pictured only in the upper layer for clarity. The two layers have different densities and there is a northward flow in the upper layer. The interface slope then represents a gradient in the x direction of density which balances the difference of horizontal velocity between the layers. The change in thickness of the layers creates gradients in PV, PV $\sim \frac{f}{h}$, such that the PV gradient vector of layer 1 is directed eastward, $\frac{\partial \mathrm{PV}_1}{\partial x} < 0$, and that of layer 2 directed westward, $\frac{\partial \mathrm{PV}_2}{\partial x} > 0$. Relative vorticity, indicated by the sense of the arrows in the vortex tubes (pictured in 1-1), is spun-up in response to the requirement that a parcel of fluid conserve its PV, while undertaking motions across an ambient PV gradient. In this model the upper and lower surfaces are rigid but the interface is flexible and may move up and down. Thus a wave in the upper layer that has induced negative vorticity, will induce the same sense of vorticity in the lower layer through a lowering of the interface. In this description growing perturbations upon the mean flow must have a phase shift in the vertical such that the relative vorticities induced by the motions in the two layers can enhance one another. This is illustrated with a 2D schematic, 1-2, which shows the displacement of the upper and lower perturbation waves with a favourable phase relation for growth. The plus and minus signs indicate

the sense of relative vorticity acquired by the fluid as it moves to this location. The effect on adjacent crests in the other layer is to amplify the motions depicted, and thus a perturbation with this phase relation between the layers will grow.

The source of energy for the growth of the perturbations is the enormous amount of energy stored in the ocean in the form of available potential energy (APE), associated with the sloping isopycnals of the large-scale flow field. This APE is an order of magnitude greater than the kinetic energy (KE) of the ocean's mean circulation [Gill et al., 1974]. Baroclinic instability provides a mechanism for extracting this APE as mesoscale eddy kinetic energy (EKE), by allowing the interface slope to flatten while giving relative vorticity to fluid parcels. The creation of these mid-ocean eddies is a first step in a process which leads to mixing and turbulence in the ocean. Thus an understanding of the conditions under which these eddies are generated, their scale and structure, has important bearings on subsequent development of ocean dynamics.

1.2 Review of relevant research

Pioneering work on baroclinic instability was done in the late 1940's by Charney[1947] and Eady[1949]. The baroclinic instability mechanism produces growing mesoscale features that are on the order of radius of deformation. In the atmosphere they are manifested in the weather's high and low pressure systems; in the ocean their scale is much smaller, on the order of 100km across. Important extension work was done by Phillips[1954], and Pedlosky[1963], and an excellent survey of baroclinic instability can be found in the text by Pedlosky[1987].

One of the more important outcomes of the early work on baroclinic instability was the derivation of a necessary criterion for baroclinic instability. The Charney-Stern (CS) criterion applied to a two-layer system [Pedlosky 1963] states that the existence of baroclinic instability in a two-layer zonal system requires a reversal of sign in the PV gradient between the layers. The presence of a planetary PV gradient, β, has a stabilizing influence since it has the same sign throughout the water column. This implies a minimum critical shear between the layers to overcome the stabilizing effect

of β, as shear induces or is associated with opposite signed PV gradients in each layer. The criteria when applied to a mid-latitude ocean, and assuming the lower layer is at rest, implies an upper layer velocity of at least 1 cm/s, necessary for the genesis of baroclinic instability. Application of this criteria to the entire ocean leaves large areas unable to produce baroclinic instability locally. Yet observations made during a cooperative experiment involving several oceanic cruises, POLYMODE(78), indicated the presence of eddy activity. It was thus supposed that radiation or advection must be important in transporting the eddies to these regions of the ocean.

In the late 1980's it was found that non-zonal flows could more easily become baroclinically unstable. This suggested the possibility that the observed eddy energy, mentioned above, could be produced locally. Pedlosky [1987] considered non-zonal flows in an unbounded domain and found that there was no critical flow speed required for instability, even the weakest flows will support growing modes. The reason is that in an unbounded environment the perturbation motions are unrestricted in their direction and can be purely zonal. As such they will not feel the planetary PV gradient, $\frac{\partial \mathrm{PV}}{\partial y} = \beta$. The zonal perturbations, $k = 0$, will have growing solutions and they will be the familiar Eady (f-plane, $\beta = 0$) solutions. With this inspiration Spall[2000] began to look at the nonlinear behaviour of non-zonal flows, with application to the instability of the interior Sverdrup flow. Both Spall[2000] and Arbic[2000]. forced non-linear models with a non-zonal component of mean shear flow and found significant eddy generation, even when the mean flow was weak. These results support the hypothesis that baroclinic instability of the relatively weak Sverdrup induced equatorward flow may be the source of large mid-ocean eddy energy. Spall proposed a scaling theory which explained the equilibrated $\frac{EKE}{KE} >> 1$ in terms of the important parameter of the basin width, L_x; that parcel motions may traverse the entire zonal extent of the basin along x-gradients of PV, acquiring large relative vorticities.

1.3 Motivation

A missing piece in the research described above is the linear analysis of a non-zonal flow in the presence of boundaries. Thus the question seemed reasonable to ask: how do variations in x, simulated by placing the flow in a channel, affect baroclinic instability of a non-zonal flow. It is natural to suppose that such boundaries in x would force motions in y that may be stabilized by planetary beta and thus some critical minimum shear for instability might be re-established.

So I have set out to find solutions of a linear stability analysis of a two layer meridional flow, sheared in the vertical, and bounded in x.

The two main questions that will be answered in this thesis are:

1/ Is there a critical shear ? If so what is it's value?
Since it is known that there exists a critical shear in the zonal case, and yet there is no such critical shear in the unbounded meridional case.

2/ What is the structure/shape of the growing perturbation?

Chapter 2

Theoretical setup and Method

2.1 Formulation

2.1.1 Setup

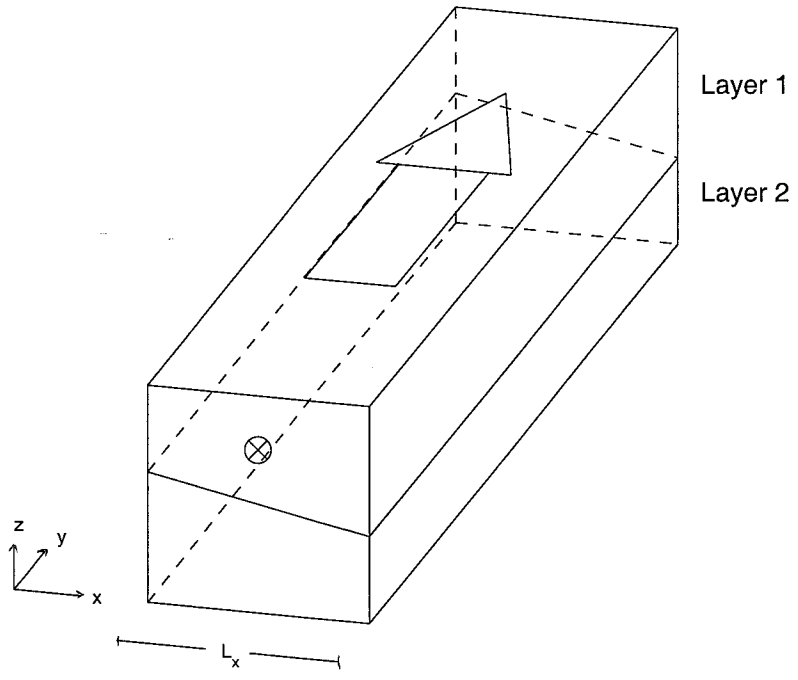

Figure 2-1: The physical setup of the problem. The channel is oriented N-S with solid boundaries at $x = 0$ and $x = L_x$. A mean flow is imposed in the upper layer only. The vertical shear of this mean flow is in thermal wind balanced with a horizontal density gradient, $\partial\rho/\partial x$. This density gradient gives rise to related gradients in PV.

We consider the simplest formulation capable of capturing baroclinic instability in a meridional channel : a two-layer, inviscid system with rigid upper and lower boundaries. The channel is infinite in the y-direction and boundary conditions of no normal flow are met at the walls, $x = 0, L_x$ (Fig. 2-1). A mean background flow is imposed in the upper layer with the other layer initially motionless. As shown in figure 2-1 the mean flow in the upper layer is of constant magnitude V and is flowing northward. The shear of the velocity in the vertical gives rise to the interface tilt in the x-direction. The slope of the interface is given by thermal wind, $\frac{\Delta z}{\Delta x} = -\frac{f}{g}\frac{V_1 - V_2}{(\rho_1 - \rho_2)/\rho_0}$, and is as drawn in figure 2-1 for poleward flow in the upper layer. The gradient of PV is opposite to the gradient of thickness and so $\frac{\partial PV}{\partial x} < 0$ in layer 1 (upper layer) and $\frac{\partial PV}{\partial x} > 0$ in layer 2 (lower layer). Both layers feel planetary beta equally; $\frac{\partial PV}{\partial y} = \beta > 0$. The oceanic relevant case of equatorward Sverdrup flow confined to the thermocline in the subtropical gyre, is a simple reflection of the + V case and will be mentioned in the summary.

2.1.2 The Equations

The framework in which to study these motions is that of quasi-geostrophy and conservation of potential vorticity. The equations for conservation of potential vorticity on a beta-plane, in each layer,

$$\begin{aligned}
\frac{D}{Dt}Q_n &= \beta_0 V_n \\
\frac{D}{Dt} &= \frac{\partial}{\partial t} + J(\Psi_n, .) \\
Q_n &= \nabla^2 \Psi_n + f_0 + \beta_0 y + (-1)^n \frac{1}{{L_{Dn}}^2}(\Psi_1 - \Psi_2) \\
\Psi_n &= V_n x + \Phi_n
\end{aligned}$$

can be combined to form the two-layer linearized quasi-geostrophic potential vorticity (QGPV) equations for the perturbation, Φ_n:

$$[\frac{\partial}{\partial t} + V\frac{\partial}{\partial y}][\nabla^2\Phi_1 + \frac{1}{{L_{D1}}^2}(\Phi_2 - \Phi_1)] + \beta_0\frac{\partial\Phi_1}{\partial x} + \frac{1}{{L_{D1}}^2}V\frac{\partial\Phi_1}{\partial y} = 0 \quad (2.1)$$

$$\frac{\partial}{\partial t}[\nabla^2\Phi_2 + \frac{1}{{L_{D2}}^2}(\Phi_1 - \Phi_2)] + \beta_0\frac{\partial\Phi_2}{\partial x} - \frac{1}{{L_{D2}}^2}V\frac{\partial\Phi_2}{\partial y} = 0 \quad (2.2)$$

Here, $L_{Dn} = \frac{\sqrt{g\frac{\Delta\rho}{\rho_0}D_n}}{f_0}$, is the deformation radius. For the equal-layer case L_D is the same for both layers. We have chosen $V_1 = V$ and $V_2 = 0$ (as stated in 2.1.1). The parameters f_0, β_0, ρ_0 and g have their usual meaning of local value of planetary vorticity, local value of latitudinal gradient of planetary vorticity, mean fluid density and gravity.

The above coupled equations can then be scaled to yield non-dimensional parameters, $F_n = \frac{L^2}{{L_{Dn}}^2}$ and $\tilde{\beta} = \frac{\beta_0 L^2}{V}$, where L is the horizontal scale of change within the system, $\frac{\partial}{\partial x} \sim \frac{\partial}{\partial y} \sim \frac{1}{L}$.

$$[\frac{\partial}{\partial t} + \frac{\partial}{\partial y}][\nabla^2\Phi_1 + F_1(\Phi_2 - \Phi_1)] + \tilde{\beta}\frac{\partial\Phi_1}{\partial x} + F_1\frac{\partial\Phi_1}{\partial y} = 0 \quad (2.3)$$

$$\frac{\partial}{\partial t}[\nabla^2\Phi_2 + F_2(\Phi_1 - \Phi_2)] + \tilde{\beta}\frac{\partial\Phi_2}{\partial x} - F_2\frac{\partial\Phi_2}{\partial y} = 0 \quad (2.4)$$

In the equal layer case $F_1 = F_2$ (as $D_1 = D_2$), and by choosing both these to have the value of 1, we allow the relative vorticity term to balance the baroclinic interaction term, so that stretching and squeezing of water columns due to motion of the interface can be compensated for with the creation of relative vorticity of the appropriate sign. This, in turn, permits exchange between the APE of the interface and eddy KE. With $\frac{L^2}{L_{Dn}^2}$ chosen to be 1, it follows that $\tilde{\beta} = \frac{\beta_0 {L_D}^2}{V}$.

In the linear case we assume a perturbation of the form

$$\Phi_n = \phi_n(x)e^{il(y-ct)}, \quad (2.5)$$

a plane wave in the y-direction with $\phi_n(x)$ being responsible for bringing the pertur-

bation amplitude to zero at the boundaries.

Substitution of the above form of the perturbation, (2.5) into equations (2.3) and (2.4) gives :

$$(1-c)[\phi_{1xx} - l^2\phi_1 - \phi_1 + \phi_2] + \frac{\tilde{\beta}}{il}\phi_{1x} + \phi_1 = 0 \tag{2.6}$$

$$-c[\phi_{2xx} - l^2\phi_2 - \phi_2 + \phi_1] + \frac{\tilde{\beta}}{il}\phi_{2x} - \phi_2 = 0 \tag{2.7}$$

Here c is normalized by the background velocity, V, and $\tilde{\beta} = \frac{\beta_0 L_D^2}{V}$, which is a measure of PV_y/PV_x. $\tilde{\beta}$ is the parameter that we are most interested in, a large value of $\tilde{\beta}$ indicates a weak shear flow.

2.1.3 The Method of Solution

At this point two routes are taken to solve for the eigenfunctions and associated eigenvalues of the linear two-layer equations. The most straight-forward route is to use the built-in MATLAB function, [c,ϕ]=eig(A), where A is the operator matrix of ϕ_n, $A\phi = c\phi$. Equations 2.6 and 2.7 can be expressed in matrix form as;

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2}{\partial x} - l^2 - 1 & 1 \\ 1 & \frac{\partial^2}{\partial x} - l^2 - 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} + \begin{bmatrix} \tilde{\beta}\frac{\partial}{\partial x} + 1 & 0 \\ 0 & \tilde{\beta}\frac{\partial}{\partial x} - 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

$$= \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} \frac{\partial^2}{\partial x} - l^2 - 1 & 1 \\ 1 & \frac{\partial^2}{\partial x} - l^2 - 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

This is possible for a major portion of the parameter space. However, the MATLAB routine became prohibitively time intensive where the growth rates are small and $\tilde{\beta}$ large, which is often concurrent with the functions being highly oscillatory in x. Where the computation time of the routine became too long, a second route is taken. This second route is a semi-analytic method:

First substitute

$$\phi_n(x) = \sum_{j=1}^{4} a_{nj} e^{ik_j x} \tag{2.8}$$

into equation (2.6) and (2.7). The elimination of the coefficients, a_n, leads to the quartic in cross-channel wavenumber,

$$\begin{aligned} k_j^4 + \left[\tfrac{(1-2c)}{c(1-c)}\tfrac{\tilde{\beta}}{l}\right] k_j^3 + \left[2(l^2+1) - \tfrac{\frac{\tilde{\beta}^2}{l^2}-1}{c(1-c)}\right] k_j^2 \\ + \left[\tfrac{(1-2c)}{c(1-c)}\tfrac{\tilde{\beta}}{l}(l^2+1)\right] k_j + \left[(l^4+2l^2) - \tfrac{l^2}{c(1-c)}\right] = 0 \end{aligned} \tag{2.9}$$

It is possible to solve (2.9) algebraically [Abramowitz, 1974], and then substitute the solutions for k into the boundary conditions:

$$\phi_n(x) = \sum_{j=1}^{4} a_{nj} e^{ik_j x} = 0 \text{ at } x = 0, L_x \tag{2.10}$$

which can be written in matrix form as:

$$\begin{array}{ll} \text{layer 1} & \text{at x = 0} \\ \text{layer 2} & \text{at x = 0} \\ \text{layer 1} & \text{at x = L}_\text{x} \\ \text{layer 2} & \text{at x = L}_\text{x} \end{array} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ e^{ik_1 L_x} & e^{ik_2 L_x} & e^{ik_3 L_x} & e^{ik_4 L_x} \\ \gamma_1 e^{ik_1 L_x} & \gamma_2 e^{ik_2 L_x} & \gamma_3 e^{ik_3 L_x} & \gamma_4 e^{ik_4 L_x} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{bmatrix} = 0, \tag{2.11}$$

where $\gamma_j = \frac{a_{2j}}{a_{1j}}$. Solutions are found where the determinate of the 4×4 matrix (2.11), generated by the boundary conditions, (2.10), is zero.

The combined results of these two methods of finding the eigenvalues are represented in figures 3-1 through 3-11 of the eigenvalue, c, as a function of along-channel wavenumber, l. This paper concludes the linear analysis by time-stepping the linear equations to further verify the existence of these growing modes. The time-stepping method used is 2nd order Runge-Kutta [Press et al., 1993].

2.2 A closer look at the equations ; initial insight

Returning to examine the equations (2.6) and (2.7) we can uncover several useful and interesting properties of the solutions apriori. These include a boundary layer structure, whose width will be different in each of the layers, and a symmetry in the eigenfunctions and eigenvalues which may spring from a symmetry of the ∇PV vectors about the along-channel axis. These properties will be graphically illuminated in chapter 3.

2.2.1 Boundary layer structure

Rewriting of equations 2.6 and 2.7 gives

$$(1-c)[\phi_{1xx} - l^2\phi_1 - \phi_1 + \phi_2] - i\frac{\tilde{\beta}}{l}\phi_{1x} + \phi_1 = 0 \qquad (2.12)$$

$$-c[\phi_{2xx} - l^2\phi_2 - \phi_2 + \phi_1] - i\frac{\tilde{\beta}}{l}\phi_{2x} - \phi_2 = 0 \qquad (2.13)$$

The presence of a zeroth, first and second derivative in x leads to the possibility of a boundary layer structure in the solution to the ODE in x. This is due to the different balances possible at different scales. The first derivative in x has an imaginary coefficient, this term will only enter the balance with other imaginary terms. Since $\tilde{\beta}$ and l are real, this leaves the eigenvalue, c, and associated eigenfunction, ϕ, as the only terms with an imaginary component. If $c_i \neq 0$ then a boundary layer structure in the function, ϕ, will result. If this c_i is positive, then by (2.5) the perturbation wave will be growing in time.

The width of the emergent boundary layer will be a function of c_i, c_r, $\tilde{\beta}$, and l. Examination of equations (2.6) and (2.7) shows that the boundary layer width will be different between the layers except in the special case of $c_r = 0.5$. This is due to the different magnitude of the factors (1-c) and $-c$. A shorthand way to explain the boundary layer behaviour is to let the second derivative balance the first derivative. From (2.13),

$$\phi_{1xx} \sim i\frac{\tilde{\beta}}{l(1-c)}\phi_{1x},$$

which implies

$$\begin{aligned}
\phi_1 &\sim e^{i\frac{\tilde{\beta}}{l(1-c)}x} \\
\phi_1 &\sim e^{i\frac{\tilde{\beta}(1-c_r+ic_i)}{l[(1-c_r)^2+c_i^2]}x} \\
\phi_1 &\sim e^{i\frac{\tilde{\beta}(1-c_r)}{l[(1-c_r)^2+c_i^2]}x}.e^{\frac{-\tilde{\beta}c_i}{l[(1-c_r)^2+c_i^2]}x}
\end{aligned}$$

similarly from (2.14)

$$\phi_2 \sim e^{-i\frac{\tilde{\beta}c_r}{l[c_r^2+c_i^2]}x}.e^{\frac{-\tilde{\beta}c_i}{l[c_r^2+c_i^2]}x} \tag{2.14}$$

The boundary layer width will depend on the magnitude of the second (decaying) exponent. A plot of the magnitude of this exponent versus wavenumber agrees well with the numerical solutions of chapter 3.

Further, there are special values of c_i and c_r that are of interest. If $c_i = 0$, non-growing solutions, there will be no boundary layer, except in the special case of $c_r = 0$ or 1 which are singular points of ϕ_2 and ϕ_1 respectively. Where $c = 0$ and 1 turn out to be important solutions and will be discussed further in section 2.2.3 and chapter 3.

2.2.2 The symmetries present in the system

This special case of equal layer depths has symmetries that deserve exploring. If we take again equations 2.6 as our starting point and set $c = \frac{1}{2} + \tilde{c}$ we obtain,

$$\begin{aligned}
\frac{1}{2}[\phi_{1xx} - l^2\phi_1 - \phi_1 + \phi_2] - \tilde{c}[\phi_{1xx} - l^2\phi_1 - \phi_1 + \phi_2] + \frac{\tilde{\beta}}{il}\phi_{1x} + \phi_1 &= 0 \\
-\frac{1}{2}[\phi_{2xx} - l^2\phi_2 - \phi_2 + \phi_1] - \tilde{c}[\phi_{2xx} - l^2\phi_2 - \phi_2 + \phi_1] + \frac{\tilde{\beta}}{il}\phi_{2x} - \phi_2 &= 0
\end{aligned}$$

Now if we let $\tilde{c} -> -\tilde{c}$ we can reobtain the above equations by setting $x -> -x$, and multiplying the equations by -1. The result is that line 1 becomes line 2 and vice versa, ie. an interchange of ϕ_1 and ϕ_2.

As an example: If $c_a = 0.42 + 0.2i$ is a root, then so is $c_b = 0.58 - 0.2i$, with $\Phi_{1a}(x) = \Phi_{2b}(-x)$ and $\Phi_{2a}(x) = \Phi_{1b}(-x)$; an East-West reflection and a switching of

layer functions between the roots. (Note: This is the same as letting $c- > 1-c$ and if c is a growing mode, 1-c is a decaying mode).

If we are interested in the growing modes,

set $\tilde{c}- > -\tilde{c}^*$

then if $\phi_1(x)$ to $\phi_2^*(x)$ and $\phi_2(x)$ to $\phi_1^*(x)$, we recover the original equations with 1 and 2 reversed.

Thus $\Phi_{1a}(y)- > \Phi_{2b}(-y)$ and $\Phi_{2a}(y)- > \Phi_{1b}(-y)$.

As an example of this case: If $c_a = 0.42+0.2i$ is a root, then so is $c_b = 0.58+0.2i$, with a relation between the eigenfunctions as expressed above; a North-South reflection and a switching of layer functions between the roots.

Thus the eigenvalues are reflected about $c_i = 0$ and also about $c_r = 0.5$. The eigenfunction symettries will be seen graphically in section 3.

2.2.3 The low-wavenumber critical point of instability

This discovery was not purely apriori, but it seems appropriate to include it here. The numerical results showed that at the low-wavenumber critical point, beyond which the neutral modes become unstable, the real phase speeds of the waves are 0 and 1, while the amplitude of the eigenfunction in one of the two layers vanishes. This is the case $c_i = 0$ and $c_r = 0$ or 1 alluded to in section 2.2.1. Acting on this information and substituting $c_r = 0, 1$ and $\phi_{2,1} = 0$ into equations 2.6 and 2.7 we reduce the equation set to a single layer equation. The single layer eigenfunction is a Rossby normal mode as is demonstrated below.

Case 1 : substitution of c=0, $\phi_2 = 0$ reduces equation 2.6 and 2.7 to a single layer equation for layer 1,

$$[\phi_{1xx} - l^2\phi_1] + \frac{1}{il}\tilde{\beta}\phi_{1x} = 0 \tag{2.15}$$

let $\phi_1(x) \sim e^{i\alpha x}$, then substitution into (2.15) yields, $-\alpha^2 + \frac{\tilde{\beta}}{l}\alpha - l^2 = 0$

solve the quadratic for α and substitute the result back into the form for $\phi_1(x)$,

$$\phi_1(x) \sim e^{i\frac{\tilde{\beta}}{2l}x} e^{\pm i\sqrt{\frac{\tilde{\beta}^2}{4l^2} - l^2}x} \tag{2.16}$$

For $\phi_1(x)$ to satisfy the BC's we require, $\frac{\tilde{\beta}^2}{4l^2} - l^2 = \frac{n^2\pi^2}{{L_x}^2}$, which can be solved for l, the low wavenumber critical point of instability,

$$l^2 = -\frac{n^2\pi^2}{2L_x^2} \pm \frac{1}{2}\sqrt{\frac{n^4\pi^4}{L_x^4} + \tilde{\beta}^2} \tag{2.17}$$

So, $\phi_1(x)$ at this location is an oscillating structure in x, $e^{i\frac{\tilde{\beta}}{2l}x}$, enveloped by $e^{\pm i\frac{n\pi}{L_x}}$, which are the Rossby normal modes of the channel.

Case 2 : substitution of c=1, $\phi_1 = 0$ reduces equations 2.6 and 2.7 to a single layer equation for layer 2. The argument follows as above except for a trivial sign change in the exponent of $\phi_2(x)$,

$$\phi_2(x) \sim e^{-i\frac{\tilde{\beta}}{2l}x} e^{\pm i\sqrt{\frac{\tilde{\beta}^2}{4l^2} - l^2}x}$$

The algebraic calculations of the low wavenumber critical point will be compared favourably to numerical result in section 3.7.

Chapter 3

Results

3.1 An overview of the results

Before launching into details of the results it may be helpful to note the key ways in which these results differ from our current knowledge of the zonal case (see App.A and [Pedlosky, 1987]), and the unbounded non-zonal case (see section 1.2 and [Pedlosky, 1987]). Simply put, the bounded meridional case supports growing modes even for very weak flows that could not go baroclinically unstable if they were zonal. In addition the growing modes are western-intensified and their horizontal structure varies throughout parameter space. The intensification of the modes in the west suggests that smaller, isolated mesoscale features will result.

At many places I will draw comparison between the zonal case and the meridional case, by this I mean simply a zonal flow in a zonal channel or a meridional flow in a meridional channel.

The figures, for the most part, have along-channel wavenumber on the abscissa. Wavenumber has an inverse relationship to wavelength, and so when I am referring to a high-wavenumber region, it can be alternatively thought of as a short-wavelength region, and vice versa.

For the purpose of section 3.2 and 3.3 a single value of non-dimensional beta, $\tilde{\beta} = \frac{\beta L_D^2}{V} = 0.5$ is used. In other places $\tilde{\beta}$ ranges from 0.1 to 1.5.

The width of the channel remains constant at a value of $L_x = 10L_D$, except for section 3.10 where variable channel widths are considered. Similarly, all figures are the results of equal layer depth calculations, except section 3.11.

3.2 A description of the eigenvalues: the phase speed, c

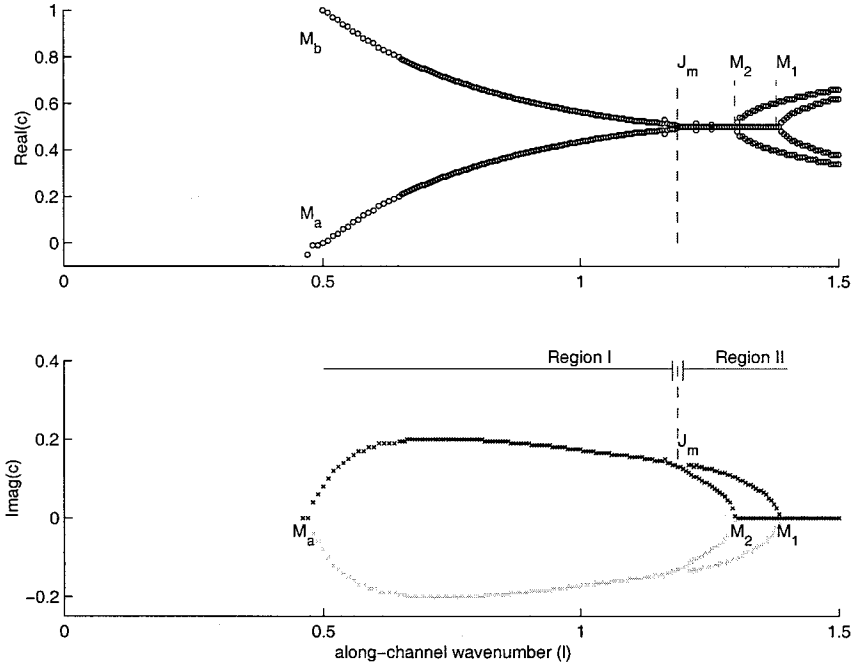

Figure 3-1: Eigenvalues, c, for $\tilde{\beta} = 0.5$ and $L_x = 10L_D$ for just the two lowest cross-channel modes. The upper panel (circles) shows the real phase speed, c_r and the lower panel (crosses) shows the imaginary phase speed, c_i, whose multiple with the along channel wavenumber, l, will give the growth rate.

In figure 3-1 the circles and crosses are the eigenvalue solutions of the two lowest cross-channel modes for $\tilde{\beta} = 0.5$. The eigenvalue curves for the zonal case for this same value of $\tilde{\beta}$ are overdrawn as thin lines for comparison in figure 3-4 and will be discussed in section 3.4. The upper panel of figure 3-1 is the real phase speed, c_r, and the lower panel is the imaginary phase speed, c_i, whose product with the wavenumber, l, will be the growth rate, ω. For each eigenvalue, c, it's complex conjugate, c^*, is also an eigenvalue, ie., the imaginary phase speed, c_i, is reflected about $c_i = 0$. We

are only interested in the growing modes, $c_i > 0$, but we will mention the structure of the decaying modes, $c_i < 0$, in section 3.9. Following along with the eye from high to low wavenumbers on figure 3-1 we first see the existence of neutral modes which join at M_1 and M_2 and become unstable. A critical point of instability is where c_i changes from being zero (not growing) to non-zero (growing). M_1 and M_2 are critical points for different growing modes.

This interval where there are two growing modes continues until the location J_m. For wavenumbers greater than that of J_m the growing modes each have, remarkably, the same phase speed of half the value of the shear velocity ($c = 0.5$). At the low end of this interval the growth rates of these two modes approach each other until they become identical. The two solutions then maintain identical growth rates but their phase speeds diverge off the $c_r = 0.5$ axis. In this figure there are always two growing modes, to the right of J_m they have the same phase speed, but different growth rates: to the left of J_m they have the same growth rate but different phase speed. These have been labelled as region II and region I in figure 3-1. Note the symmetry of the real phase speed about the $c_r = 0.5$ axis (this was referred to in section 2.2.2). At low wavenumbers there is a second critical point where c_i vanishes. This is the longwave stability threshold, denoted here as M_a, M_b, where $c_r = 0$ and 1. These points were discussed in section 2.2.3. Beyond this wavenumber, it is difficult to trace the stable solution, but these are not our prime interest.

Figure 3-2 shows the eigenvalues for the first six cross-channel modes for $\tilde{\beta} = 0.5$ including the first two that were shown in figure 3-1. The neutral (stable) curves that were included for illustration in figure 3-1 are not redrawn here. The symmetry in real phase speed about $c_r = 0.5$ persists, and the same apparent pairing of modes continues. Although only six are shown, there are an infinite number of unstable modes present in the meridional case. Squashed into the low-wavenumber end of the spectrum are modes with cross-channel structure that vary rapidly in x, and these modes have vanishingly small growth rates.

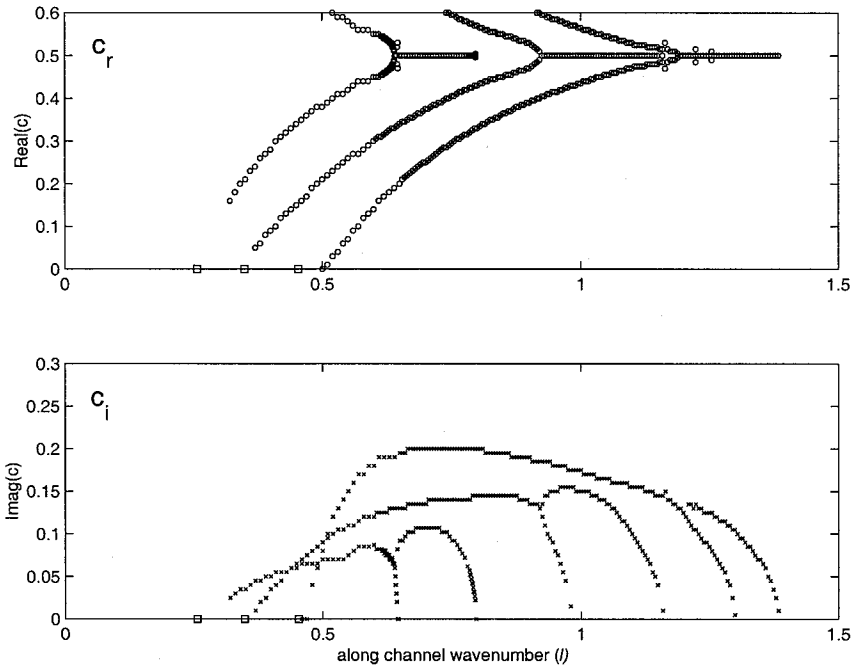

Figure 3-2: Eigenvalues, c, for $\tilde{\beta} = 0.5$ and $L_x = 10L_D$. Unstable eigenvalue solutions with higher cross-channel structure are drawn. There are, in fact, an infinite number of paired unstable modes present, but not drawn. They exist squashed into the low-wavenumber region. The boxes plotted in the lower panel mark the low-wavenumber critical point of instability for the first three cross-channel Rossby normal modes (see sections 2.2.3 and 3.6).

3.3 A first look at the structure of the eigenfunctions

Figure 3-3 has figure 3-1 inset in the top left and is surrounded by a series of eigenfunctions for various points along the curve. Each eigenfunction is labelled with a letter and has both a left and a right panel, corresponding to the upper and lower eigenfunction. Here we see that unlike the zonal case, the shape of the mode alters throughout parameter space. In the zonal case the modes retain their sinusoidal structure throughout wavenumber-space, varying only in the respective amplitude between layers. In the meridional case the eigenfunctions alter between layers and from wavenumber to wavenumber. Again starting the description at high wavenumbers: A1 and A2 show the neutral modes with very short along-channel wavelength. Each is a wave with amplitude, $|\phi_n|$, in one layer approaching 0 in the limit of the real phase speeds approaching 0 and 1 at large wavenumbers. This is the same limit as the zonal case. At B these two waves have interacted and become a growing mode and already there is a slight westward intensification and a slight vertical phase shift. A second growing mode is represented by the branch on which C can be found. This mode has a higher cross-channel structure. In this region, to the right of D, labelled as region II in 3-1, the growing modes have c_r at 0.5.

The symmetry between the layers of B and C is a special case of the second example of section 2.2.2, where $\tilde{c_r} = 0$ leads to one layer being a reflection of the other layer about the cross-channel axis, (a N-S reflection), thus the maximum amplitude of each layer is the same, and the BL width is also the same. At D there is a double root: two identical solutions. At E we can see that the above-mentioned symmetry is expressed in the more general sense where $\tilde{c_r} \neq 0$ and so the N-S reflection and exchange of layers is between E1 and E2. These are now two growing modes with the same growth rate, but with phase speeds reflected about the $c_r = 0.5$ axis. A similar reflection and exchange between layers for F, G and H would give the appropriate eigenfunction for the $c_r > 0.5$ branch. Recall also the first example of symmetry mentioned in 2.2.2, where the growing and decaying mode are related by an E-W

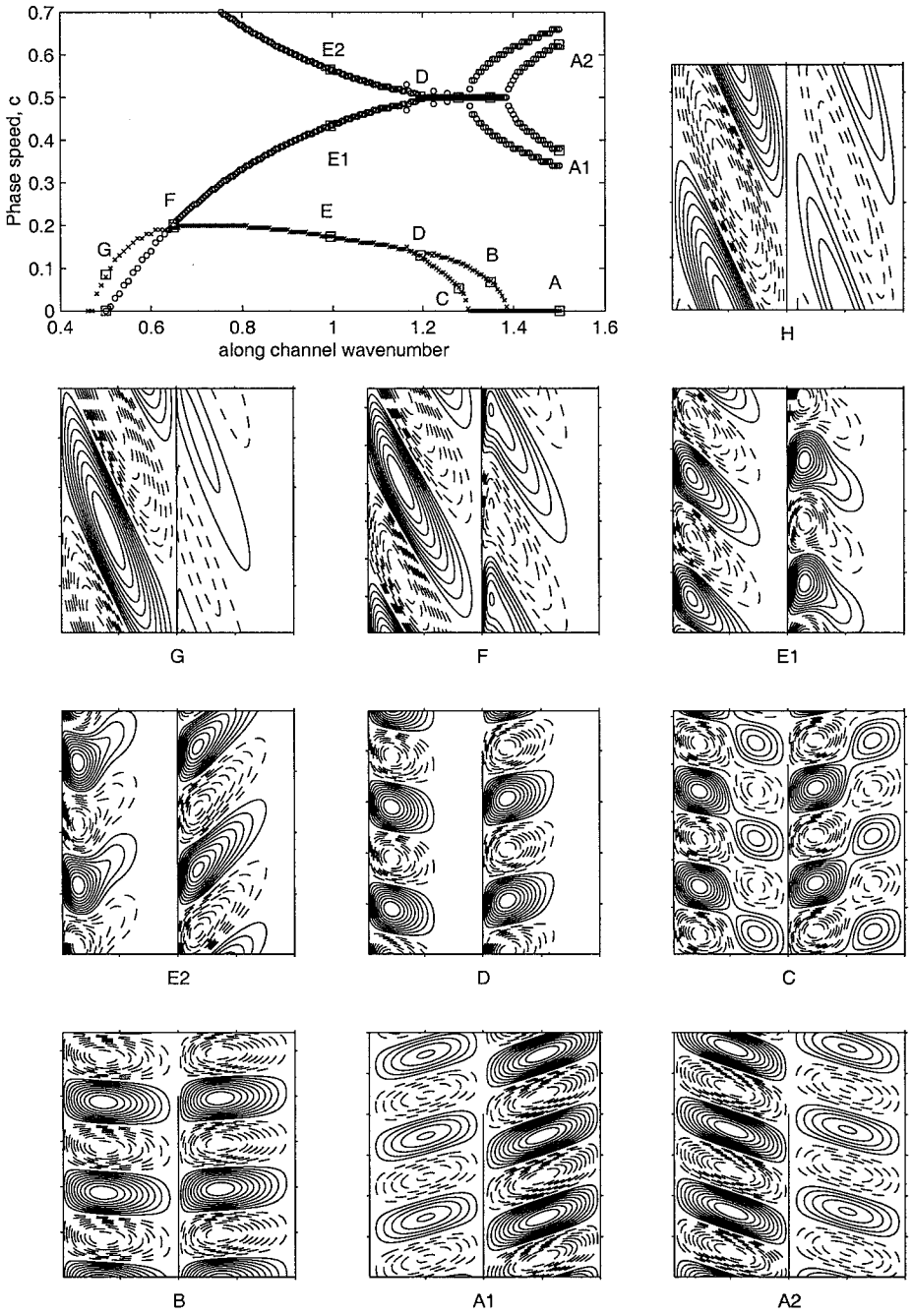

Figure 3-3: The eigenvalues, c, (top left) and eigenfunctions, Φ_n, for $\tilde{\beta} = 0.5$ and $L_x = 10L_D$. Each eigenfunction is labelled with a letter and is split into two panels, the left being the upper layer eigenfunction, Φ_1, and the right the lower layer eigenfunction, Φ_2. The eigenfunctions have been compacted in x, their true proportions can be attained by stretching the eigenplots in the x-direction by a factor of 2. The phase speed and growth rate can be found by reading off the letter from the eigenvalue curve in the top left. A1, A2 and H are neutral modes, and H exists outside the axis of the above plot.

reflection and exchange of layers. Thus the solutions that are decaying, $c_i < 0$, are eastern intensified. By location F the amplitudes of the eigenfunction have become quite disparate between the layers. As we move toward G the growing eigenfunction has a vanishing amplitude in the lower layer, this vanishing eigenfunction is also highly intensified in the west. The difference in boundary layer structure, as was alluded to in 2.2.1, is apparent here as the western intensification falls off in the upper layer, while it gains strength in the lower layer as c_i and c_r move to 0. Where both c_i and c_r are 0, at $l = .465$, Φ_2 has zero amplitude and Φ_1 is a Rossby normal mode. An analysis of this point was included in section 2.2.3. Location H, with $l = 0.45$ is in the neutral region. This is a stable wave with a phase speed of $c_r = -3.1$ and so is not plotted on the eigenvalue curve as it is outside the axis range. Both layers are in phase, and the eigenfunction becomes more barotropic in the limit as c_r rapidly approaches negative infinity along this branch toward decreasing wavenumber.

For the region of instability represented here, $l = 0.465 - 1.38$, the fastest growing perturbation comprises not just one, but two modes. Both E1 and E2 have the maximum growth rate, $\omega_{max} = (c_i l)_{max}$, in the spectrum, but they are propagating at different phase speeds. This is an interesting phenomenon of linear amplitude vacillation, and has important bearing on non-linear wave-wave interactions, but is a curiosity specific only to the equal-layer case.

3.4 Comparing the meridional solutions to the zonal solutions

Figure 3-4 is a comparison plot for the zonal and meridional case with $\tilde{\beta} = 0.5$ and $L_x = 10L_D$. The thin lines are the two lowest cross-channel modes of the zonal case. The crosses and circles are the eigenvalues for the meridional case as we have already seen in figure 3-1.

The two zonal modes remain independent and retain their sinusoidal structure throughout wavenumber-space, varying only in the respective amplitude between lay-

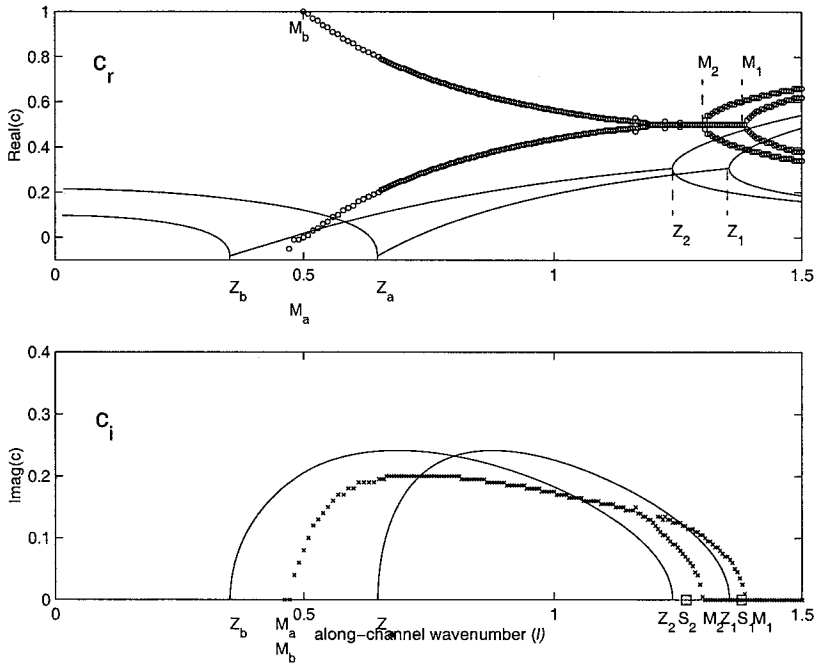

Figure 3-4: Eigenvalues, c, for $\tilde{\beta} = 0.5$ and $L_x = 10L_D$. The crosses and circles are the eigenvalues for the meridional channel for the two lowest cross-channel modes (as in figure 3-1). The solid lines are the eigenvalue curves for a zonal channel of the same width, for the two lowest cross-channel modes (see also App A).

ers. $\phi(x) \sim sin(\frac{mode\pi}{L_x})$ for mode $= 1, 2$. What is so curious about the meridional case is that what appear to start out as independent modes at the high wavenumber end, are actually co-dependant modes at the low wavenumber end. In fact, as will be examined in section 3.8, all the modes are interdependent.

At the high-wavenumber end we can draw some similarity between the zonal and meridional cases. The squares on figure 3-4 indicate the wavenumber of critical instability for $\tilde{\beta} = 0$ for mode 1 and 2. These are included here and are important because at $\tilde{\beta} = 0$ the directional bias vanishes, and the meridional and zonal case are identical. Z1 and Z2 are critical points for modes 1 and 2 of the zonal case and M1 and M2 are critical points for the meridional case as was mentioned in section 3.2. At the locations M1, Z1, the eigenfunction have a similar structure, indeed as $\tilde{\beta}$ is decreased to 0, M1 and Z1 converge upon the square, S1. Similarly M2 and Z2 converge upon S2.

An at first surprising finding illustrated here is that as $\tilde{\beta}$ increases from 0 the

critical wavenumber decreases for the zonal case but increases for the meridional case. This will be discussed in section 3.7.

For the zonal case the unstable modes have phase speeds less than 0.5 and their low wavenumber critical points have phase speed just less than 0. For the meridional case the high wavenumber critical points have phase speeds of exactly 0.5, the eigenvalues are symmetric about $cr = 0.5$ and their low wavenumber critical points have phase speed of 0 or 1.

Another way in which the solutions of the meridional case differ from the zonal case is in the number of unstable modes that are present in the system. For the channel width of $10L_D$ there are only four unstable modes in the zonal case whereas there are an infinite number of unstable modes in the meridional channel of the same width.

3.5 Dependence on $\tilde{\beta}$

Figure 3-5 illustrates how the structure of the eigenvalues for the lowest cross-channel modes evolve for increasing value of $\tilde{\beta}$. Beside each eigenvalue curve is drawn the fastest growing mode that would emerge from a white noise perturbation at this value of $\tilde{\beta}$, labelled with it's growth rate. Remember, that increasing $\tilde{\beta}$ is equivalent to decreasing the shear of the velocity.

As $\tilde{\beta}$ increases the growth rate ($c_i l$) decreases, but remains non-zero for $\tilde{\beta} > 1$, while the fastest growing mode occurs at ever increasing along-channel wavenumber. As $\tilde{\beta}$ increases the high wavenumber critical point moves to larger values as does the low wavenumber critical point, and consequently all the modes are shifted to shorter wavelengths. For the last four values of $\tilde{\beta}$ there are two fastest growing modes as was illustrated by E1 and E2 in figure 3-3. For $\tilde{\beta} < 1$ the fastest growing mode is always the one with the lowest cross-channel structure. This is not guaranteed for $\tilde{\beta} > 1$, although it does, in this case, hold true for $\tilde{\beta} = 1.1$. In the case of $\tilde{\beta} = 1.5$ the fastest growing mode has a more rapidly oscillating cross-channel structure, and this will be discussed in more detail in section 3.8 and figure 3-6.

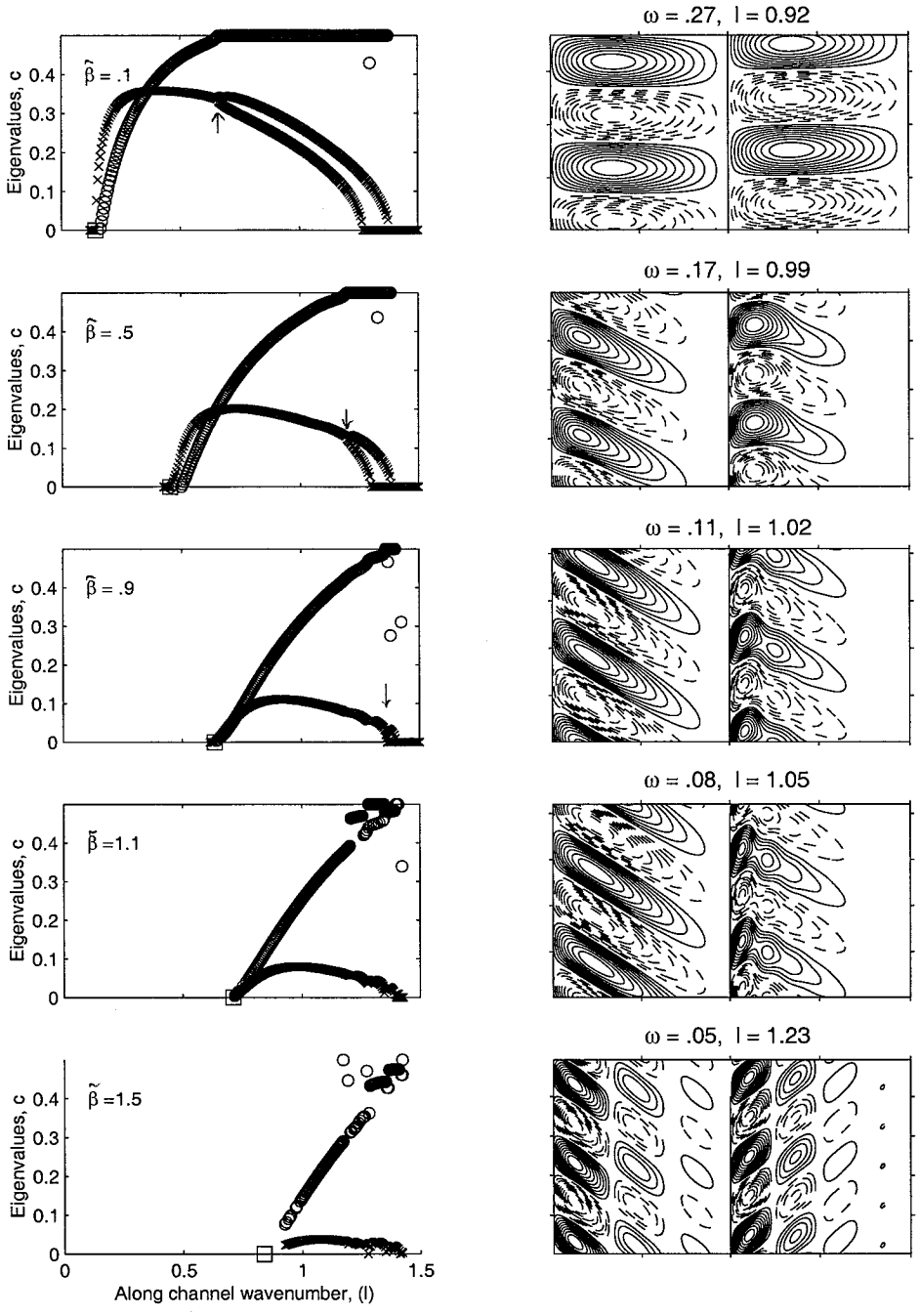

Figure 3-5: Eigenvalue curves for varying values of $\tilde{\beta}$ are shown in the left column. $\tilde{\beta} = 0.1,\ 0.5,\ 0.9,\ 1.1,\ 1.5$. In the right column are the eigenfunctions of the fastest growing mode for the respective values of $\tilde{\beta}$. Their growth rate and along-channel wavenumber are printed above each eigenfunction. A box plotted in each of the left-hand panels marks the low-wavenumber critical point of instability.

As $\tilde{\beta}$ increases from 0.1 to 0.9 we see that the wavenumber that separates region I from region II, indicated by the arrows, also increases. What happens to this point as $\tilde{\beta}$ becomes > 1 ? As will be shown in section 3.8, mode 1 and 2 do not remain together until the low wavenumber cutoff. There is a succession of unstable joining modes, and the critical point where they join does not continue to have $c_r = .5$.

As $\tilde{\beta}$ increases we also see that the boundary layer becomes sharper, more western intensified, as $\tilde{\beta}$ is a parameter in the exponent of equation (2.14).

3.6 The low-wavenumber limit

On each eigenvalue curve in figure 3-5 a box denotes the low-wavenumber critical point where $c_i = 0$. The actual boxes placed on the plot are calculated algebraically, by substitution of n = 1 into eqn. 2.17 of section 2.2.3, and are thus independent of the numerical results. The mode number, n, is chosen to be 1 for the lowest across-channel mode. The agreement between the algebraic and numerical calculations suggests that it is indeed a Rossby normal mode in just one layer that is becoming unstable at the low wavenumber cutoff. Agreement is also seen in figure 3-2 where the wavenumbers of the first three Rossby normal modes, n=1,2,3 of eqn. 2.17, are calculated and plotted as boxes.

It is hard to determine what the stable part of the curve looks like for wavenumbers less than this because $c = 0, 1$ are singular points of the equations. One part of the stable eigenvalue curve has been traced for wavenumbers less than the singular points. This mode is H (fig. 3-3), as mentioned in 3.3, and it's image, related in the same way as E1 and E2.

3.7 The high-wavenumber cutoff

The critical point where the lowest mode becomes unstable at the high-wavenumber end is also called the short-wavelength cutoff. Returning to figure 3-4 and the discussion of 3.4, the boxes on panel two indicate critical points for $\tilde{\beta} = 0$, S1 and S2. One

sees by comparing the meridional case, M1 and M2, to the boxes of $\tilde{\beta} = 0$ that the high wavenumber cutoff actually increases, as $\tilde{\beta}$ increases ie. shorter along-channel wavelengths become unstable. This finding that l_{crit} increased as $\tilde{\beta}$ increased was reinforced by a perturbation analysis about $\tilde{\beta} = 0$ (Pedlosky, 2000, personal communication). Values of the high-wavenumber cutoff of instability have been calculated for $\tilde{\beta} = 1, 5, 100$ as $l = 1.39, 2.35, 10$. This seems somewhat odd, as it states that as the magnitude of the shear decreases to weak velocities the region in wavenumber-space that supports unstable modes expands, although granted the growth rates are much smaller.

3.8 $\tilde{\beta} > 1$

For the zonal case $\hat{\beta} = \frac{\beta_0 L_D^2}{U} = 1$ is the critical value beyond which there are no unstable growing modes. This cutoff does not exist in the meridional case and there are unstable modes for $\tilde{\beta} = \frac{\beta_0 L_D^2}{V} > 1$ as was seen in section 3.5. This may be because even when $\tilde{\beta} > 1$ there are regions where the PV gradient is of opposite sign between the two layers, which ceases to be the case at $\hat{\beta} = 1$. This difference will be discussed in detail in the section on PV vectors (4.1)

There is a qualitative change in the character of the eigenvalues and companion eigenfunctions for $\tilde{\beta} > 1$. Figure 3-6 is an enlarged plot of a subregion of the last panel in figure 3-5. The figure shows only non-dimensional wavenumbers in the range $l = 1.2 - 1.5$ for $\tilde{\beta} = 1.5$. The two upper-left panels of figure 3-6 are the eigenvalues, c_r and c_i. The lettered points correspond to the eigenfunctions of the same letter. The numbered points do not have their eigenfunctions plotted but are mentioned in the description that follows.

For $\tilde{\beta} > 1$, at the high-wavenumber (short-wavelength) end, neutral modes coalesce to become unstable for narrow wavenumber intervals. The progression in wavenumber space shows two modes joining for periods of instability and then separating as neutral waves to join with a different mode and again become unstable. This process may be repeated a number of times. As an example, we describe the progres-

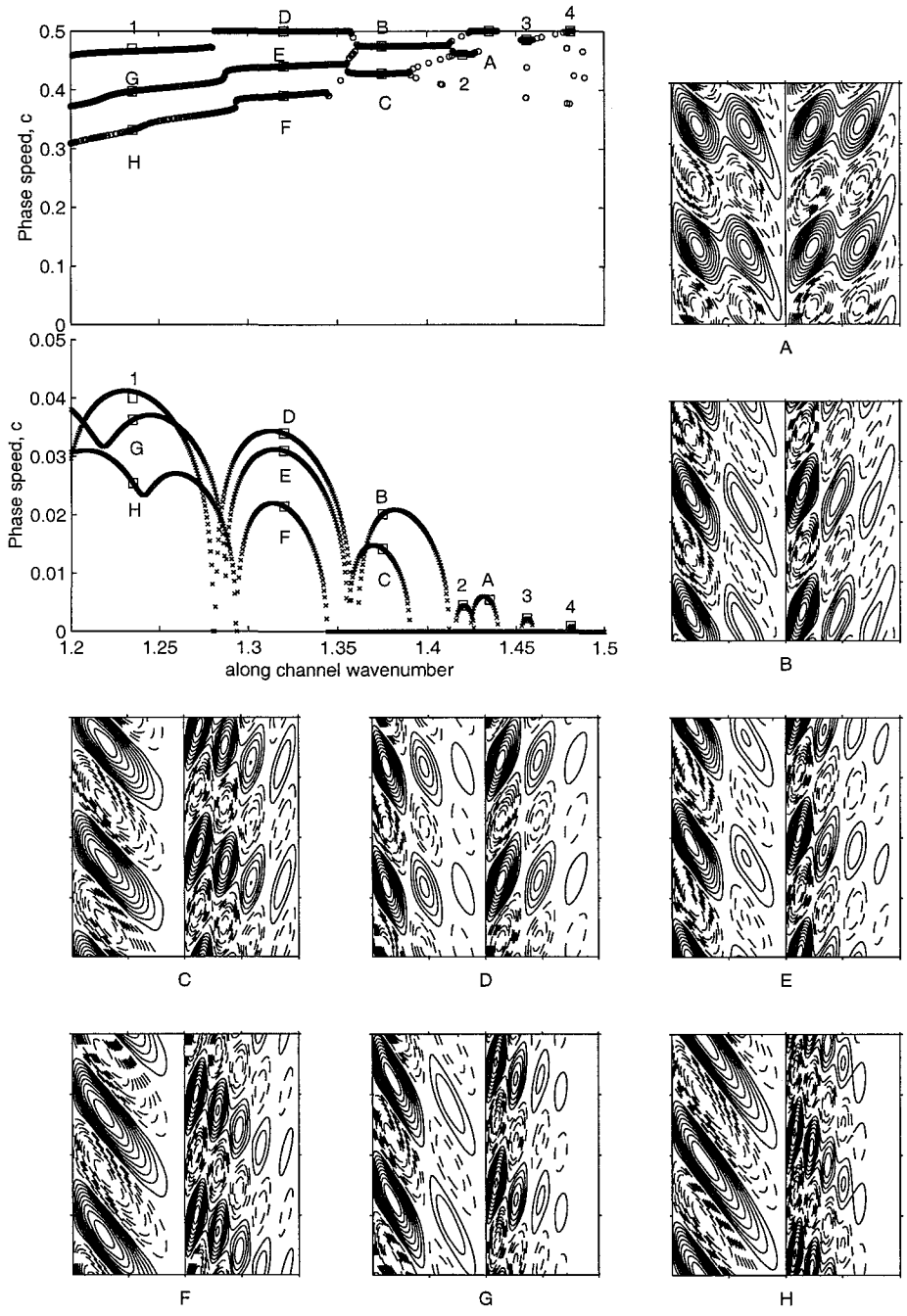

Figure 3-6: As an example of $\tilde{\beta} > 1$: Eigenvalues and selected eigenfunctions in the high wavenumber region for $\tilde{\beta} = 1.5$ and $L_x = 10L_D$.

sion of a mode from point 4 to 3 to 2 to C to F to H. Throughout the transition this mode maintains an upper-layer structure with a single across-channel maximum while the lower layer structure becomes progressively more oscillatory in the x-direction. The growing mode, point 4, is formed by the joining of two neutral modes with a single across-channel structure and itself has a single across channel structure in both layers, similar to point B in figure 3-3, and it's phase speed is $c_r = 0.5$. Where this mode becomes stable again the two neutral modes branch off the $c_r = 0.5$ axis. The exiting neutral mode with $c_r < 0.5$ intersects with the neutral mode of double across channel maxima before it can reach $c_r = 0.5$, and so becomes a growing mode, point 3, with a single across channel structure in the upper layer and a double across channel structure in the lower layer. The same progression occurs for points 2, C, F, and H, as neutral modes with progressively higher cross channel structure become unstable. The unstable mode H has a After H the mode does not become stable again until the low-wavenumber cutoff at $l = 0.84$ denoted as the box in figure 3-5 in panel 5. The progression described above appears again when narrow channel widths are considered in section 3.10.

The point labelled 1 is the fastest growing mode for $\tilde{\beta} = 1.5$ and the eigenfunction was drawn in figure 3-5.

Results for $\tilde{\beta} > 1.5$ are not shown in this thesis, simply because the growth rates are so very small, and the intervals of wavenumber for which individual modes are unstable, so very short.

3.9 Flux Calculations : Heat, Vorticity, and Momentum

In figure 3-7 the various fluxes are shown for the fastest growing mode for $\tilde{\beta} = 0.5$. The heat flux, $\overline{u\theta}$, as expected, is negative, indicating that the heat flux is westward, such that the result would be a flattening of the interface, and a loss of potential energy from the system. The heat flux, like the eigenfunction, is intensified in the

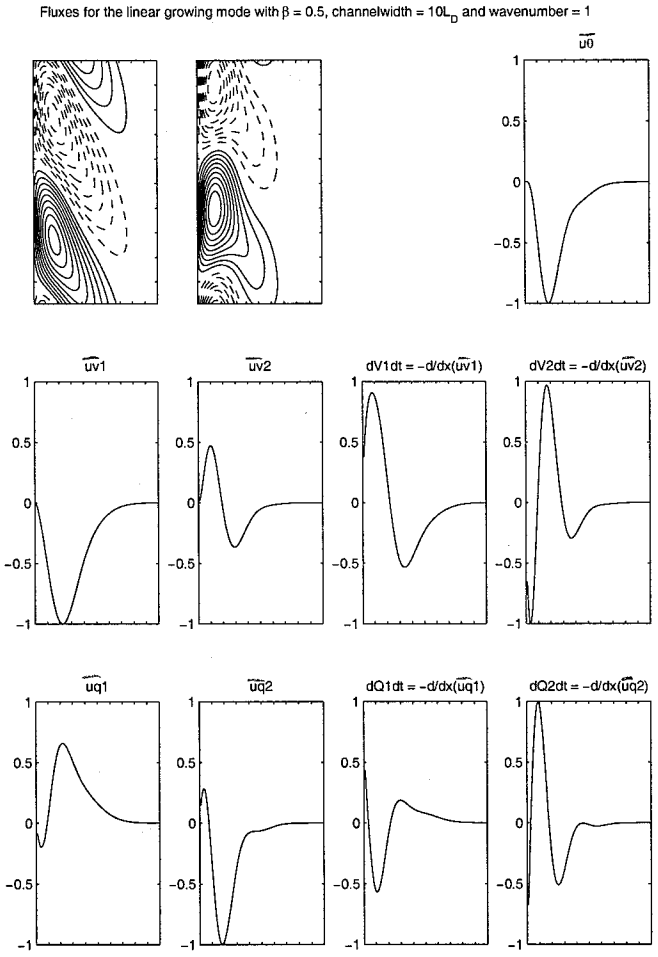

Figure 3-7: The two layer eigenfunction under consideration is shown in the upper left. The panels show heat flux, $\overline{u\theta}$, momentum flux, $\overline{uv}$, and vorticity flux, $\overline{uq}$. The right panels show how the convergence of the flux might affect the mean fields of velocity and vorticity, V and Q, in a non-linear system.

west.

The vorticity flux, $\overline{uq}$, is also western intensified, as is the momentum flux, $\overline{uv}$, which is notably absent, or zero, in the zonal case. Figure 3-7 are the fluxes for just one of the two modes with the same growth rate, the momentum and vorticity fluxes for the other mode are related by: $\overline{uv_{1a}} = -\overline{uv_{2b}}$ etc. Thus the sum effect of the two modes gives a momentum flux that is equal in magnitude but opposite in sign for the two layers, similarly for the vorticity flux. The convergence of the perturbation vorticity flux leads to changes in the mean vorticity field, as indicated in the last two panels of figure 3-7. The ambient vorticity, Q, can then be inverted to find the mean flow, V.

3.10 Channel width

Unlike the zonal case, there appears to be no lower limit on channel width for instability. For the zonal case, the minimum channel width to have just one mode unstable, is $L_{Xmin} = \frac{\pi}{\sqrt{2}} \sim 2.2L_D$. For $L_X < L_{Xmin}$ the zonal flow, no matter how strong, will always be stable.

In the meridional case a very thin channel still supports growing modes, although their growth rate is much decreased. As the channel width decreases, the high wavenumber cutoff of instability also decreases. In the case where the channel width $= 0.5L_D$, as shown in the first panel of figure 3-8, the growing modes have intervals of instability being almost confined to a single wavenumber.

Figure 3-8 shows eigenvalue solutions, c, for four different channel widths: $\frac{L_x}{L_D} =$ 0.5, 2, 10, and 50. Panel 3 is the one with which we are familiar, with $L_x = 10L_D$. The upper part of each panel is the real phase speed, c_r. In these it is easy to see that the coincidence of two neutral modes brings about instability. For $L_x = 0.5L_D$ instability occurs just at these crossover locations, and higher resolution in wavenumber is needed to find all of them. For the case $L_x = 2L_D$ there is an interval of wavenumber, $l = 0.35$ to 0.52, where mode 1 is unstable with $c_r = 0.5$ either side of this it is neutral. At $l = 0.3$ it meets the neutral mode 2 curve and becomes again unstable, much like the

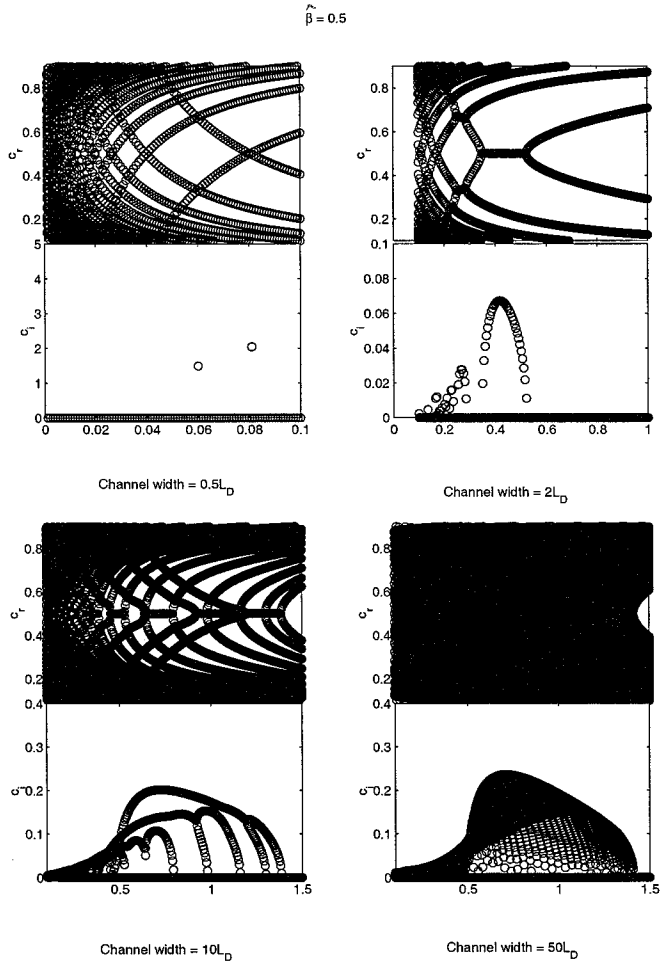

Figure 3-8: The four panels show eigenvalues, c, with $\tilde{\beta} = 0.5$ for different channel widths: $\frac{L_x}{L_D} = 0.5,\ 2,\ 10,\ 50$. In each case the upper panel is c_r and the lower is c_i.

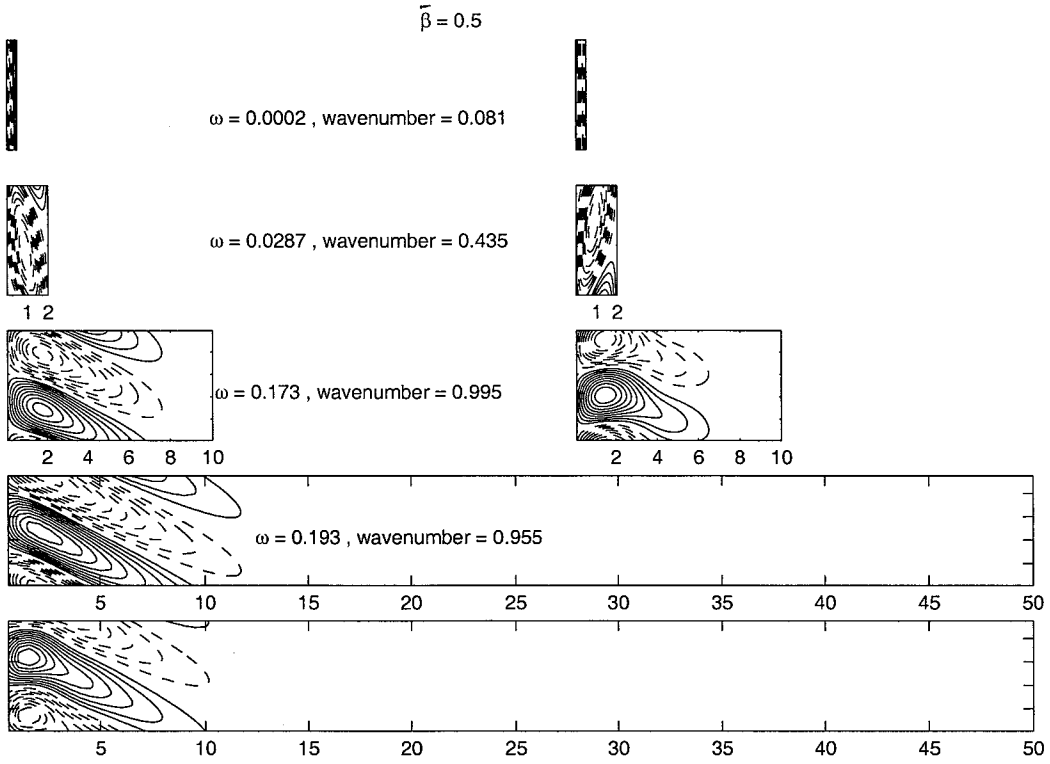

Figure 3-9: The fastest growing eigenmodes for the variable channel widths of figure 3-8. The x and y are approximately to scale.

high wavenumber region of $\tilde{\beta} = 1.5$, as discussed in section 3.8. For $L_x = 10L_D$ we have the scenario already discussed in earlier sections and there is no real difference in character for $L_x = 50L_D$ excepting that far more modes are unstable at higher wavenumbers.

Figure 3-9 shows the fastest growing eigenfunctions for the channel widths of $\frac{L_x}{L_D} = 0.5,\ 2,\ 10$, and 50. The channel widths are shown to scale, and the x-y scale is very near correct so the slope as seen can be taken as realistic. The slope apparent in the eigenfunction alters little once the channel width gets beyond $10L_D$. The growth rates and wavenumber of the fastest growing modes are marked adjacent to the figure. As the channel broadens the maximum growth rate increases. The unbounded Eady growth rate would be 0.29 at this wavenumber.

3.11 Unequal layer depths

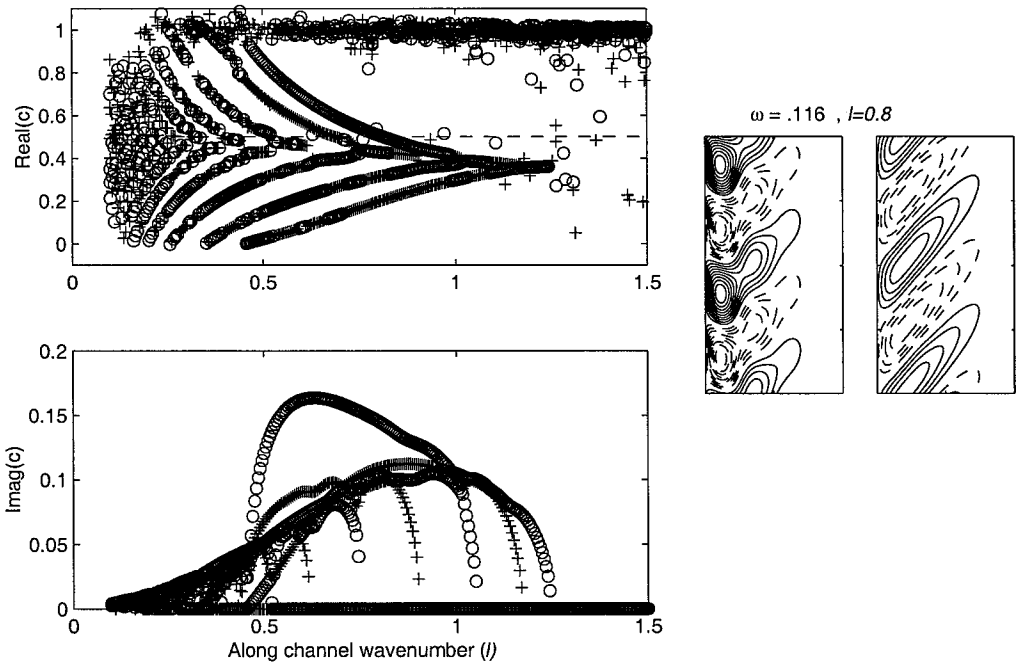

Figure 3-10: Unequal layer depths: $F_{1,2} = [8/5\ 2/5]$ and $\tilde{\beta} = 0.5$. The lower layer is four times the thickness of the upper layer.

The depth ratio, $\frac{D_1}{D_2}$ is chosen to be $\frac{1}{4}$ which is representative of the thermocline and deep ocean. The case of unequal layer depths breaks the symmetry referred to in

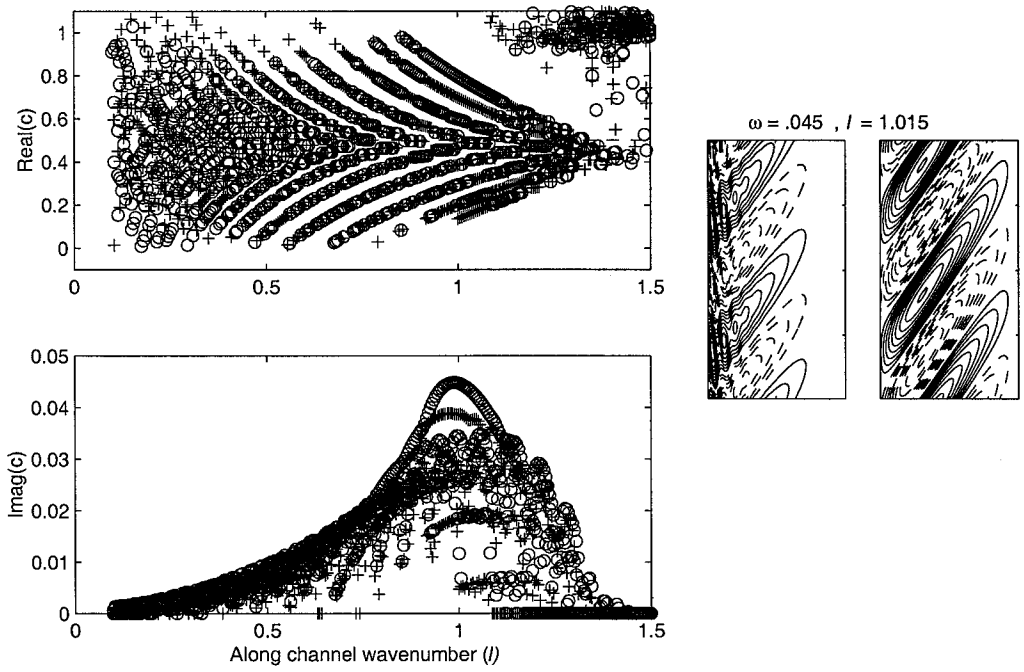

Figure 3-11: Unequal layer depths: $F_{1,2} = [8/5\ 2/5]$ and $\tilde{\beta} = 1.5$. The lower layer is four times the thickness of the upper layer.

section 2.2.2 and there is now only one fastest growing mode. Figure 3-10 and 3-11 show the eigenvalues, c, and the fastest growing eigenfunction for $\tilde{\beta} = 0.5$ and 1.5.

Comparison with figure 3-5 shows that the growth rates are not significantly different. However the along-channel wavenumber of the fastest growing mode has decreased from $l = 0.99$ to $l = 0.8$ in the case of $\tilde{\beta} = 0.5$ and from $l = 1.23$ to $l = 1.015$ in the case of $\tilde{\beta} = 1.5$. The high wavenumber cutoff has also decreased. The fastest growing mode now favours a wave which exists on the branch $c_r > 0.5$ with phase speed approaching the upper layer velocity. This is why the eigenfunctions have the opposite slope to those in figure 3-5, which represented the $c_r < 0.5$ branch of the symmetric eigenvalues. These fastest growing modes pictured in 3-10 and 3-11 would very clearly result in tight eddy structures of the initial instability, especially in the upper layer.

If V were negative, a southward flow in the upper layer, the fastest growing mode would still have an intense boundary layer structure in upper layer. The difference between the case with +V and -V is the slope of the eigenfunction in the horizontal

plane, instead of a NE-SW slope the fastest growing mode would have a NW-SE slope.

A property that is not altered by the change in layer depths is the low wavenumber critical point, as the equations of section 2.2.3 are unaffected by the value of F.

Although there are no mirror image eigenvalue and eigenfunction pairs, the modes are effected by the existence of each other. This can be seen by close examination of figure 3-10 at $l = 0.65$. Where a new mode becomes unstable, an existing growing mode deviates from a smooth eigenvalue path, both the c_r and c_i curves have a kink here.

3.12 Time Evolution

3.12.1 Verification

To verify the emergence of these modes the coupled linear equations were time-stepped using a second order Runge-Kutta time-stepping. Finite differencing was used in x with $\Delta x = 0.1 L_D$, and $\frac{\partial}{\partial y} = il$, The initial perturbation in all cases was a sine curve. The results give growth rates that match the theory. Figure 3-12 shows the growth, on the left, and eigenfunctions, on the right, for both equal ($\tilde{\beta} = 0.5, 1.5$ $F_{1,2} = [1\ 1]$) and unequal layers ($\tilde{\beta} = 1.5$ $F_{1,2} = [8/5\ 2/5]$). The solid line plots the maximum amplitude as a function of time, the dashed line is the predicted growth rate. In each case the eigenfunction shown on the right is the fastest growing mode (or modes). In the first two cases, two modes are growing at the fastest growth rate, and the resultant eigenfunction is the sum of these traveling through each other at their different phase speeds. This explains the oscillation of the solid line about the dashed line, a phenomenon coined as linear vacillation. The period of the oscillation is a function of the difference of their phase speeds. The last case of unequal layers has just one mode growing at the fastest rate.

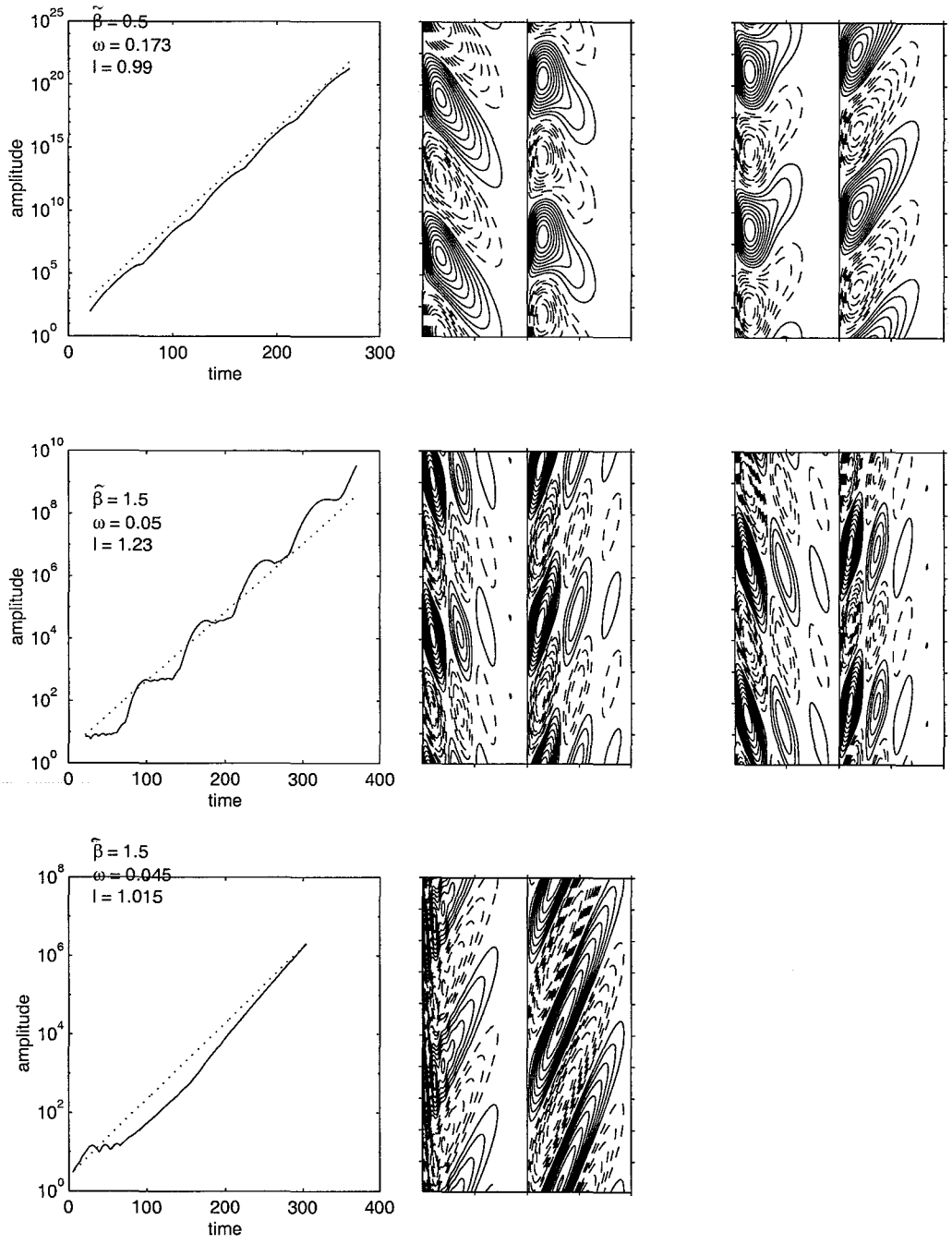

Figure 3-12: The time evolution of a random perturbation at fixed parameter values. The left column plots the growth of the amplitude. The right column shows the fastest growing eigenfunction(s). The first two rows are with equal layer depths, $\tilde{\beta} = 0.5$ and 1.5 respectively, and the third row is with unequal layers, F=[8/5 2/5], and $\tilde{\beta} = 1.5$.

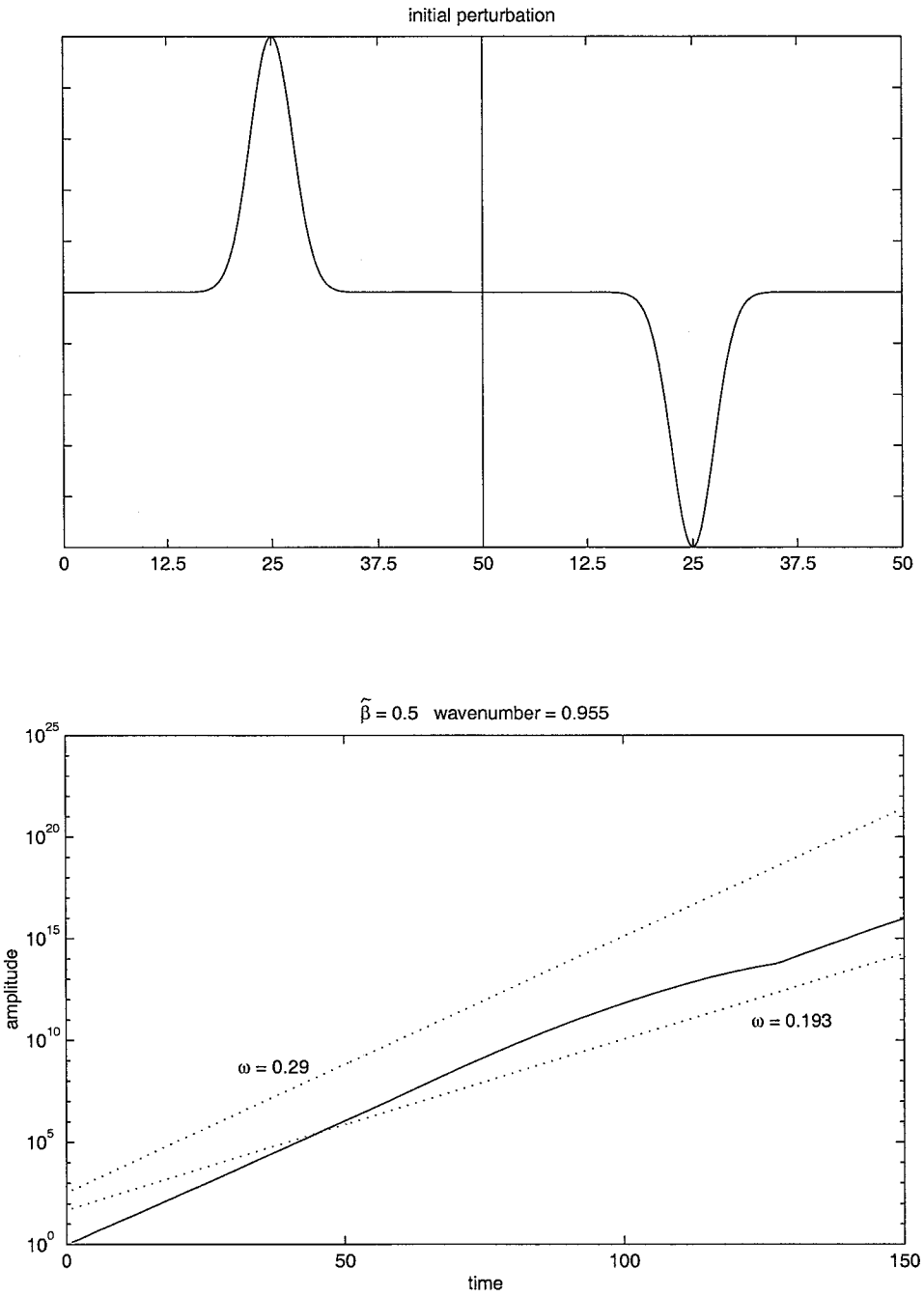

Figure 3-13: The time-evolution of an initial perturbation that is far from the boundaries. $\tilde{\beta} = 0.5$, wavenumber $= 0.955$ and $L_x = 50L_D$. The upper panel is a cross-channel slice through the initial perturbation; a gaussian of opposite sign in the two layers. The solid line in the lower panel is the growth of the perturbation. The dashed lines are the growth rate at this wavenumber for an Eady mode, $\omega = 0.29$, and the western intensified meridional mode,$\omega = 0.193$.

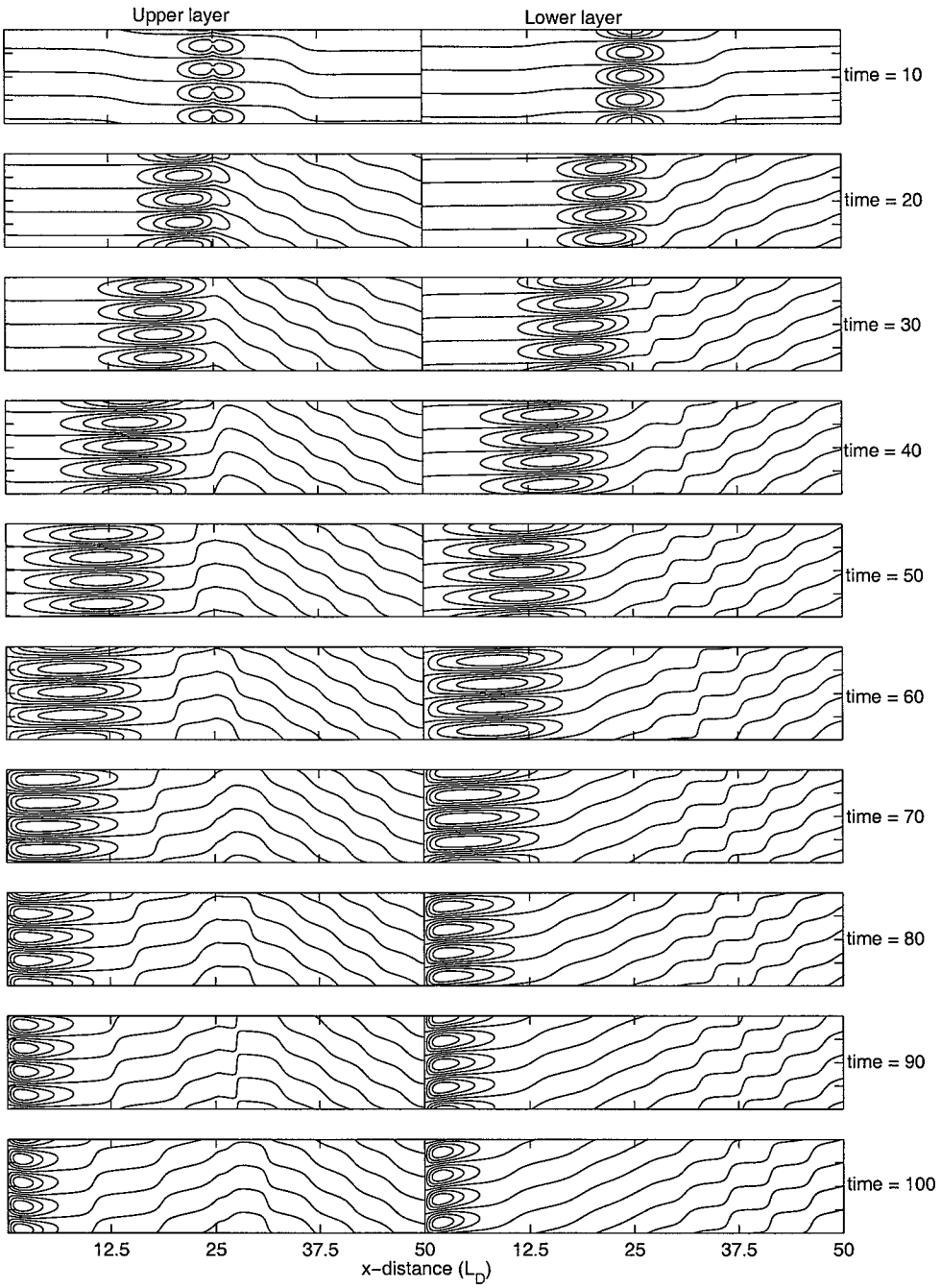

Figure 3-14: The time-evolution of an initial perturbation that is far from the boundaries. $\tilde{\beta} = 0.5$, wavenumber $= 0.955$ and $L_x = 50L_D$.

3.12.2 Wide Channel Limit

It is supposed that in the wide channel limit, a perturbation, if confined to the centre of the channel, would not feel the effect of the boundaries, and may initially evolve as if it were unbounded. This has in fact been observed by Spall[2000] and so in an attempt to reproduce these result with a linear evolution of such a system I have time-stepped a gaussian perturbation centred within a channel of width $L_x = 50L_D$. The time-stepping method is the same as above with $\Delta x = 0.25L_D$. Figure 3-13 shows the initial perturbation function in the upper panel and the time evolution of the amplitude in the lower panel. Snapshots spanning the first 100 timesteps are shown in figure 3-14. The along channel wavelength is chosen to be 0.955. Initially the perturbation moves as a packet to the west, elongating slightly until it reaches the boundary. During the transit time to the western boundary the perturbation is growing at the Eady growth rate of 0.29, as shown in figure 3-13. There follows a transition period as the growth rate slows down to 0.193, the calculated growth rate of the meridional bounded mode, and subsequently the evolution resembles the first panel of figure 3-12 as the two growing modes vascillate.

Chapter 4

Discussion and Summary

4.1 Some understanding from the ∇PV vectors.

A physically intuitive way of explaining baroclinic instability was described in 1.1 with the aid of figure 1-2. For the induced relative vorticity to cause growth of the perturbation it was shown that the gradients of ambient PV in the two layers needed to be of opposite sign. It is then easy to see why the critical value of $\hat{\beta} = 1$ exists for the zonal case. For $\hat{\beta} < 1$ the ∇PV vectors in the two layers are directed opposingly, the angle between them is (180^o). Fluid motion along any trajectory will experience a negative ∇PV in one layer and a positive ∇PV in the other, and indeed for all $\hat{\beta} < 1$ there are growing modes. However for $\hat{\beta} > 1$ the ∇PV vectors are in the same direction, thus a fluid in motion will feel either an increase in ambient vorticity in both layers or a decrease in both layers. For this situation no growing modes exist.

The meridional case is quite different. Since both x and y components of the ambient PV gradient exist, the angle between the ∇PV vectors in the two layers spans a range between 0^o and 180^o (Figure 4-1). This angle decreases as $\tilde{\beta}$ increases, but never vanishes. So there will always be a range of trajectories along which fluid parcels will feel an increase in ambient PV in one layer while feeling a decrease in the other, thus explaining why weak flows, $\tilde{\beta} > 1$ can support eddy growth. The angle γ between the ∇PV vectors can be expressed in terms of $\tilde{\beta}$, $\gamma = 2\tan^{-1}\frac{1}{\tilde{\beta}}$. The arc, γ, sketched on figure 4-1 represents the angle of motion available where ∇PV_1 and

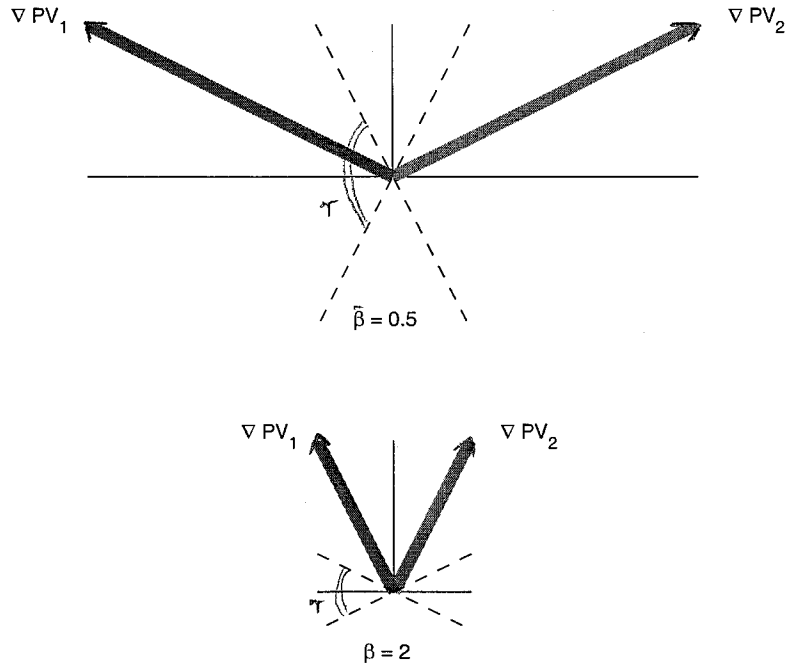

Figure 4-1: The ∇PV vectors of the two-layer system. The solid lines show the direction of the gradient of PV in the two layers. The dashed lines show the direction of no PV change in each layer. Upper and lower layer are indicated by the subscripts 1 and 2. The vector directions encompassed by γ are the regions where the PV gradients in the two layers have opposite sign and so motions in these directions will induce relative vorticity of fluid parcels in the two layers of the opposite sign.

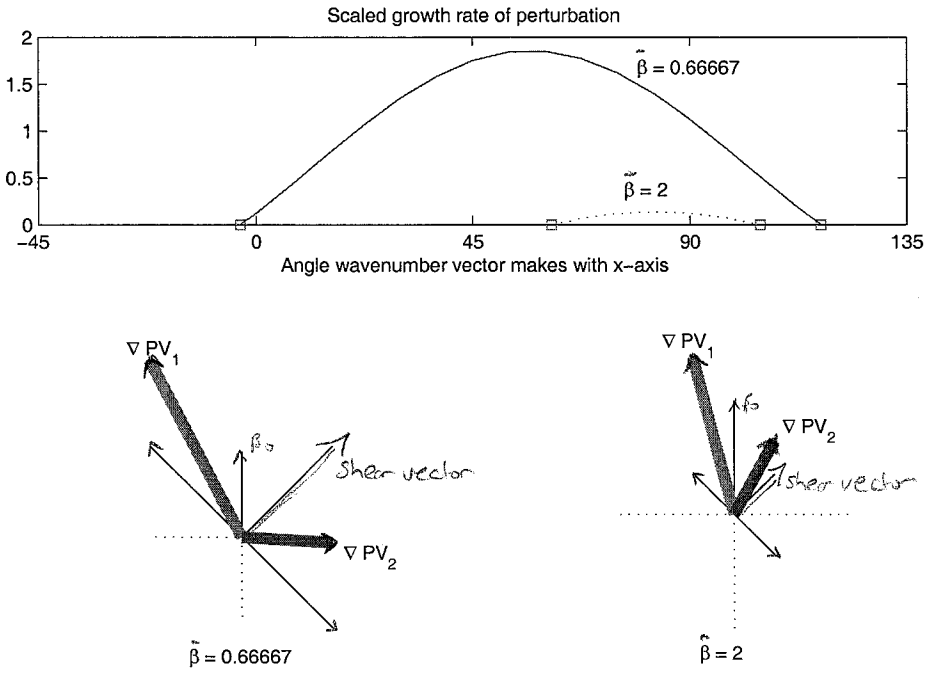

Figure 4-2: Upper panel : growth rate of wave in an unbounded environment as a function of angle of wavenumber vector. The mean shear flow is at an angle of 45^o. Lower panels : schematics of the components of ∇PV for two different values of $\tilde{\beta}$, $\tilde{\beta} = 2/3$ and 2.

∇PV_2 along the direction of motion are of opposite sign. Fluid parcels moving along trajectories within the arc of γ will pick up opposite signed relative vorticity in the two layers and so create the possibility of growth if the phase relation between the layers is favourable (see figure 1-2).

This requirement that the fluid motion is along paths where the PV gradient of one layers is positive and the other negative can be demonstrated in figures 4-2 and 4-3. The upper panel shows the growth rate as a function of the angle of the wavenumber vector for an unbounded shear flow. The shear flow can have an arbitrary angle and the two cases shown here are 45^0 in figure 4-2 (a NE shear flow) and 90^0 in figure 4-3 (a N shear flow). The upper panel of figure 4-2 is in a sense a reproduction of a result calculated by Pedlosky from equation 7.13.14 of his 1987 text. The lower panels are schematics of the PV gradient vectors of the layers. The boxes plotted on the upper panel indicate the angles ∇PV_1 and ∇PV_2 make with the x-axis. The boxes coincide with the critical points of instability. What this states is that the wavevector of a

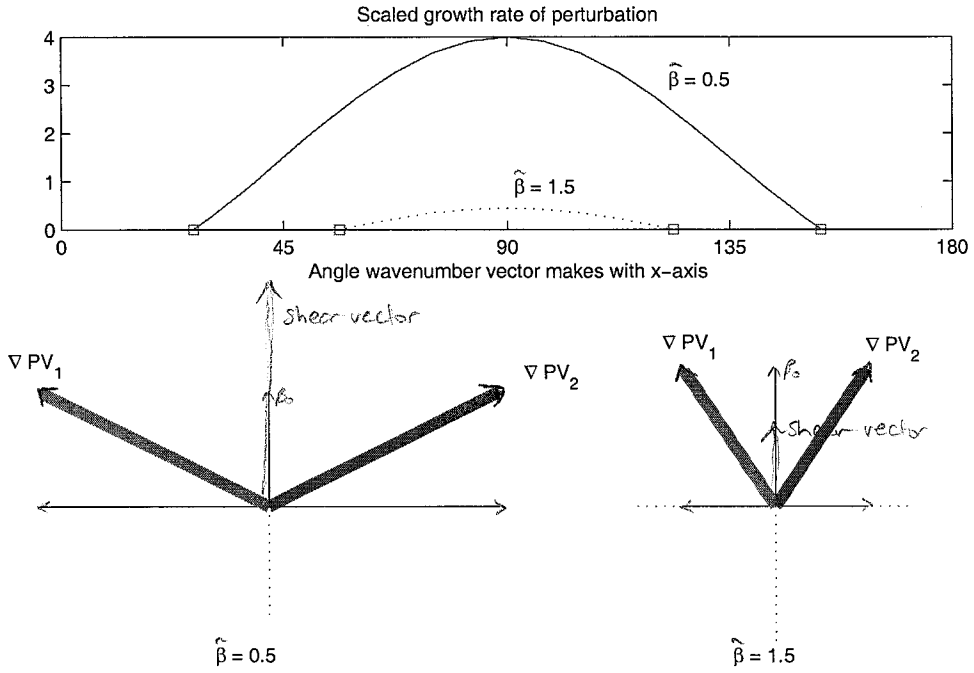

Figure 4-3: Upper panel : growth rate of wave in an unbounded environment as a function of angle of wavenumber vector. The mean shear flow is at an angle of 90^o. Lower panels : schematics of the components of ∇PV for two different values of $\tilde{\beta}$, $\tilde{\beta} = 1/2$ and $3/2$.

wave that is growing exists only within the angle between the PV gradients of the two layers. Consequently, the direction of motion, transverse to the wavevector, must exist within the angle γ as represented in 4-1 for growth to occur.

The slope in the horizontal plane of the eigenfunctions pictured in the results is such that fluid trajectories for the majority of the channel extent see a change in sign in gradient of PV between the layers. Where the general slope goes beyond that of the region γ the growth rate drops of dramatically. In addition the phase relation as seen in the growing modes has layer 2 leading layer 1 which is the discretization of the phase tilting westward, against the shear, in the zonal Eady growing modes.

4.2 The results and discussion section

I have presented the eigenvalue and eigenfunction solutions of the linear, two layer QGPV equations in a meridional channel with a mean flow imposed in the N-S direc-

tion in the upper layer (see the setup of fig. 2-1). A meridional component of mean flow leads to PV gradient in x in addition to the PV gradient in y, β. In the equal layer case this arrangement leads to the ∇PV vectors in the two layers being mirror images of each other about the y-axis (fig. 4-1), and this can be thought of as the source of the symmetric relation of the eigenfunctions and eigenvalues about $c_r = 0.5$. The angle between these two vectors, γ, is a function of the relative strengths of the ambient PV gradients in the coordinate directions of x and y; $\mathrm{PV_y} = \beta_0$, and $\mathrm{PV_x} = \mathrm{V/L_D^2}$. The parameter $\tilde{\beta}$ is a measure of this ratio, $\tilde{\beta} = \frac{\beta_0 L_D^2}{V}$. It was found that for all values of $\tilde{\beta}$ there are growing modes present, and it is suggested that the reason behind this is simply that there remains a direction, denoted by γ, where the ∇PV vectors in the two layers are of opposite sign and so growth may occur as described in section 1.1.

The growing eigenfunctions are found to have a boundary layer behaviour which is manifested as intensified recirculations in the west. The width of the boundary layer scales as $\frac{\beta c_i}{l[c_r^2+c_i^2]}$. The eigenvalues are reflected about $c_i = 0$ and $c_r = 0.5$. There are an infinite number of unstable modes progressing into the long wavelengths, but with rather weak growth rates.

In the equal layer case the symmetry leads to an apparent coupling of modes. At low wavenumbers, region I of figure 3-1, there are two modes with the same growth rate, whose structures are related by a reflection about the x-axis and a exchange between layers. At a point, which I denoted as J_m, there is a change in regimes. For wavenumbers higher than J_m, region II, the two modes now have different growth rates, and maintain a constant phase speed of $c_r = 0.5$. The symmetry is now restricted to a N-S reflection between layers within an individual mode. The two growing modes now have an unrelated structure (as c_i is different) and begin to take on the character of a cross-channel oscillation, $k = \frac{n\pi}{L_x}$ as c_i approaches 0.

For $\tilde{\beta} < 1$ the modes are paired : 1&2, 3&4, 5&6, etc.. (fig. 3-2). For $\tilde{\beta} > 1$ it may be mode 1&4 that pair, and in separating, as we move to higher wavenumbers, 4 remains neutral while 1 joins with 3, and then 2, and then 1 (fig. 3-6). Because of this the fastest growing mode is not necessarily the one with lowest cross-channel structure (see 3-5) and is likely to become very complicated, and with short wavelengths as $\tilde{\beta}$

increases.

As $\tilde{\beta}$ increases (decreasing shear values), the growth rate decreases, but the wavenumber of maximum growth increases, and the high-wavenumber cutoff increases.

I have attempted to explain why these non-zonal flows become unstable at such small speeds in terms of the requirement of opposing signs of ∇PV between the layers. However there is no similar proof as can be obtained in the zonal case.

4.3 The concise summary

One of the more important questions of recent oceanographic interest concerns the mixing of the ocean. This thesis examines the genesis of baroclinic instability which is the first step in a process which ultimately leads to turbulence and mixing.

It is known that unbounded non-zonal flows are unstable to all shear strengths, and nonlinear work has shown that indeed weak nonzonal flows evolve into large eddy variability. What we didn't know was the specifics of the genesis of such instability, particularly in an environment where the background has variations in x.

Our first important finding is that even when the background flow, on which the instability grows, is bounded in extent, it remains that all flow strengths are unstable. While we thought planetary beta might stabilize the flow by requiring a critical shear, this was not the case and was explained by the existence of the angle between ∇PV_1 and ∇PV_2 for all values of shear. To the contrary, planetary beta could be said to have a destabilizing effect as it allows much shorter wavelengths to become unstable.

The second major finding is that the structure of the growing modes are western intensified, because of planetary beta. Therefore the growing perturbations exhibit tight recirculations in the west at the boundary. However a wave packet isolated from the boundaries will evolve as an Eady mode, without noticeable intensification until it makes contact with the boundary. Thus for a growing perturbation to exhibit the boundary layer structure of chapter 3 before becoming nonlinear the perturbations need to be near the western boundary or of a scale comparable to the distance between boundaries. Thus we expect only basin size modes to exhibit the boundary layer structure before the growing perturbations become non-linear.

What we can now state is that perturbations will grow locally in a weak non-zonal shear environment, that the initial perturbations, if able to feel the western boundary, have recirculations in the western part of the domain and that the presence of boundaries does not critically stabilize a non-zonal channel flow. That, in fact, the entire ocean is subject to the genesis of baroclinic instability is the principal result of our work.

Appendix A

Zonal case summary

For ease of reference I have outlined here the results of a linear instability analysis of the two-layer system with a zonal mean flow sheared in the vertical. The results can be found in detail in the text of Pedlosky [1987] and are only briefly described below.

The two-layer linearized quasi-geostrophic potential vorticity (QGPV) equations for the perturbation, Φ_n, are: (compare with eqns. 2.1 and 2.2)

$$[\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}][\nabla^2\Phi_1 + (\Phi_2 - \Phi_1)] + (\beta + U)\frac{\partial \Phi_1}{\partial x} = 0 \quad \text{(A.1)}$$

$$\frac{\partial}{\partial t}[\nabla^2\Phi_2 + (\Phi_1 - \Phi_2)] + (\beta - U)\frac{\partial \Phi_2}{\partial x} = 0 \quad \text{(A.2)}$$

substitution of the perturbation waveform, $\Phi_n = \phi_n(y)e^{ik(x-ct)}$, division by Uik and scaling of $L \sim L_D$ yields (compare with eqns. 2.6 and 2.7)

$$(1-c)[\phi_{1xx} - k^2\phi_1 - \phi_1 + \phi_2] + (\hat{\beta} + 1)\phi_1 = 0 \quad \text{(A.3)}$$

$$-c[\phi_{2xx} - k^2\phi_2 - \phi_2 + \phi_1] + (\hat{\beta} - 1)\phi_2 = 0 \quad \text{(A.4)}$$

with $\hat{\beta} = \beta_0 L_D^2/U$. By setting $\phi_n(y) = \sum_{j=1}^{4} a_{jn}e^{il_j(y-ct)}$ substitution and elimination of either a_1 or a_2 leads to an algebraic relation which can be solved for the eigenvalue, c, or the cross-channel wavenumber, l.

The equation for c leads to

$$c = \frac{1}{2} - \frac{\hat{\beta}(K^2+1)}{K^2(K^2+2)} \pm \frac{1}{2}\frac{\sqrt{K^4(K^4-4)+4\hat{\beta}^2}}{K^2(K^2+2)} \tag{A.5}$$

which reduces to the Eady solution, $(c = \frac{1}{2} \pm \sqrt{\frac{K^2-2}{K^2+2}})$ for $\hat{\beta} = 0$. $K^2 = k^2 + l^2$.

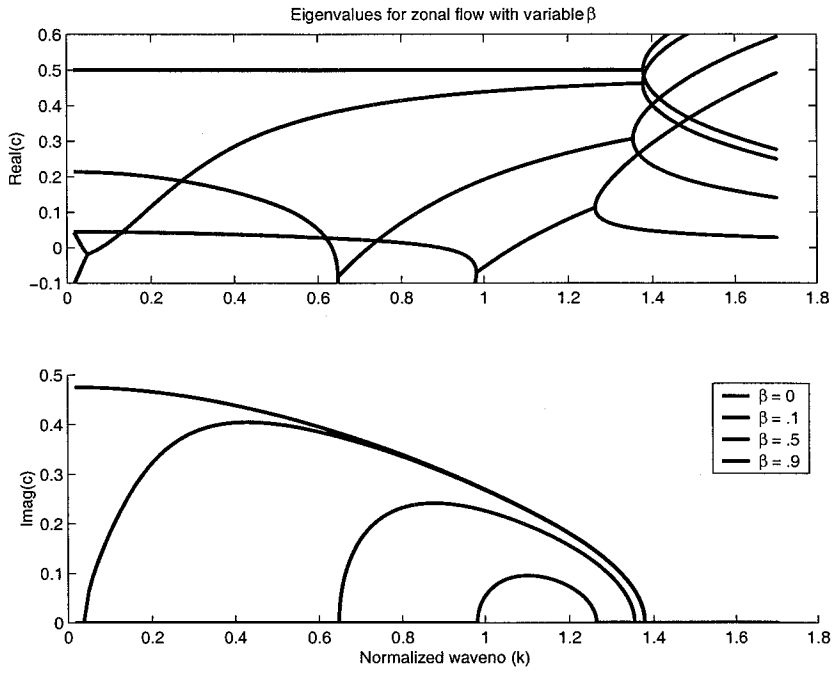

Figure A-1: Eigenvalues, c, for a zonal channel for $\hat{\beta} = 0,\ 0.5,\ 0.9$ and $L_y = 10L_D$ for the lowest cross-channel modes.

Figure A-1 shows the eigenvalue curves, c_r and c_i, for $\hat{\beta} = 0,\ 0.5,\ 0.9$ for the lowest cross-channel mode. Anywhere along the curves the eigenfunction has a simple sinusoidal structure in the cross-channel direction, ie. $l_j = n\pi/L$ where $n = 1$.

Equation A.5 yields two important criteria for instability: that $\hat{\beta} < 1$ and that $K^2 < 2$. The first comes from requiring that the radicand be negative, so $\hat{\beta}^2 < \frac{K^4(4-K^4)}{4}$. The RHS of this inequality has a maximum value of 1 at $K^2 = \sqrt{2}$ value of 1 at $K^2 = \sqrt{2}$ which implies that $\hat{\beta}_{max} = 1$. The second comes from setting $\hat{\beta} = 0$ and asking that the radicand be negative.

Appendix B

Numerical proceedure

MATLAB code for both methods of solving for the eigenvalues and eigenfunctions are available.

The numerical proceedure of the second method is represented in words below.

Choose a value of $\tilde{\beta}$ and L_x.

Make a matrix of c within the apriori expected range of $[c_r = -.1 : 0.6, c_i = -.1 : 0.5]$.

Now loop through wavenumber looking for solutions.

For $l = 0 : l_{max}$ (I chose $l_{max} = 1.6 > \sqrt{2}$ which is the apriori expected value)

. Solve for $k's$, $k(\tilde{\beta},\ l,\ c)$ using eqn. 2.9

. If value of determinate = 0 (from eqn. 2.12) a solution to the BC's

. then solve for $a's$ and

. record the eigenvalues, c and eigenfunctons, Φ as solutions

. end

end

The advantage of this method is that the $k's$ and $a's$ are calculated as intermediate steps and examination of the character of the $k's$ and $a's$ can help us better understand the character of the solutions. Evaluating the coefficients, a_j, leads to the discovery that only three of the four are non-zero (note: in the zonal case only two are non-zero - which results in the simple sinusoidal structure). The zero coefficient multiplies the exponent with the k that leads to intensification in the east.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, New York, 1972.

[2] B. K. Arbic. *Generation of mid-ocean eddies: The local baroclinic instability hypothesis.* PhD thesis, MIT-WHOI Joint Program in Physical Oceanography, 2000.

[3] J. G. Charney. The dynamics of long waves in a baroclinic westerly current. *J. Meteorology*, 4:135–163, 1947.

[4] B. Cushman-Roisin. *Introduction to Geophsical Fluid Dynamics.* Prentice Hall, 1994.

[5] E. T. Eady. Long waves and cyclone waves. *Tellus*, 1(3):33–52, 1949.

[6] W. H. Press et al. *Numerical Recipes in C.* Cambridge University Press, 1993.

[7] A. E. Gill, J. S. A. Green, and A. J. Simmons. Energy partition inthe large-scale ocean circulation and the production of mid-ocean eddies. *Deep Sea Res.*, 21:499–528, 1974.

[8] J. Pedlosky. Baroclininc instability in two-layer systems. *Tellus*, 15:20–25, 1973.

[9] J. Pedlosky. The stability of currents in the atmospere and ocean, part i. *J. Atmospheric Science*, 27:201–219, 1974.

[10] J. Pedlosky. *Geophysical Fluid Dynamics*, chapter 7. Springer Verlag, New York, second edition, 1987.

[11] N. A. Phillips. Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus*, 6:273–286, 1954.

[12] POLYMODE(78). Poly mid-ocean dynamics experiment.

[13] M. Spall. Mechanism for low-frequency variability and salt flux in the meditteranean salt tongue. *J. Geophysical Research*, 99:10121–10129, 1994.

[14] M. Spall. Generation of strong mesoscale eddies by weak oceangyres. *J. Marine Research*, 58:97–116, 2000.